Visualization of Medicare data

Introduction:

Medicare, a cornerstone of American healthcare, is a federal health insurance program primarily designed for people who are 65 years of age or older, certain younger individuals with disabilities, and those suffering from End-Stage Renal Disease (ESRD) which is permanent kidney failure requiring dialysis or a transplant. This program plays a pivotal role in providing healthcare access to a significant segment of the U.S. population, thereby impacting numerous aspects of public health and policy.

The motivation for our project, titled "Visualization of Medicare Data," stems from the ongoing changes and challenges within the Medicare system, especially in light of demographic shifts and economic pressures. With an extensive dataset sourced from data.gov covering the years 2013 through 2022, our study seeks to harness data visualization techniques to offer new insights into the Medicare landscape. These insights aim to aid policymakers, healthcare providers, and beneficiaries in understanding and navigating the complexities of Medicare.

Research Questions:

Our project is guided by the following critical questions, which address both the operational and strategic aspects of Medicare:

1. **Enrollment and Service Utilization Trends:** How have Medicare enrollments and service utilization changed over the years, particularly in response to demographic shifts and policy changes?
2. **Drug Prescription Dynamics:** What are the patterns in drug prescriptions among Medicare beneficiaries, and how do these reflect on the program's healthcare needs and cost implications?
3. **Physician Services and Geographic Distribution:** How do the availability and type of Medicare-registered physicians vary across different states, and what does this mean for access to healthcare services among Medicare beneficiaries?

By exploring these questions, our project seeks to provide a comprehensive view of the Medicare program's operational dynamics, identify areas for improvement, and predict future trends in healthcare needs and provisions under Medicare.

Description of data:

To conduct a comprehensive analysis of Medicare, we utilized several datasets taken from data.gov, covering a range of aspects related to Medicare beneficiaries, prescription drugs, and physician services. The datasets span various years and provide a granular view of the Medicare landscape. Below is an overview of the datasets employed in our study:

1. **Medicare Enrollment Data (2013 - 2023) - medicare_from_2013_to_2023:**
   Contents: This dataset covers Medicare enrollments across various Medicare plans (Parts A & B, D), including demographic information such as age groups, disability status, and ESRD status. It comprises 26 columns and 470323 records. It tracks changes and trends in enrollments for about a decade, offering insights into the shifting landscape of Medicare beneficiaries.
2. **Medicare Drugs Data (2017 - 2021) - drugs_data:**
   Contents: The dataset provides detailed information on prescription drugs under Medicare from 2017 to 2021, including brand and generic names, manufacturers, total spending by year, total dosage units, and total claims. It also includes average spending per dosage unit, claim, and beneficiary, highlighting trends in drug usage and costs over the five years. This dataset comprises 46 columns and records for 13,751 drugs, reflecting the wide variety of medications covered under Medicare. Key variables include yearly spending, claims, dosage units, and beneficiary counts.
3. **Medicare Physicians and Medical Equipment Data - physicians_data:**
   Contents: This dataset includes data on physicians who treat Medicare beneficiaries, detailing their specialties, distribution across states, practice types, and RUCA (rural-urban commuting area codes: urban vs. rural). It also covers information on medical equipment provided under Medicare, looking at rental vs. purchase patterns and the costs covered by Medicare for these equipment. However, it does not specify whether the data was for the year 2021 or until 2021.

The datasets are no longer available in their original links submitted during the project proposal. However alternative links are provided above. The analysis proceeded with previously downloaded files.

Data Dictionary:

Medicare Enrollment Data		
Variable Name	**Term description**	**Definition**
Year	Year	Indicates the calendar year of Medicare enrollment
Month	Month	Indicates the month of Medicare enrollment
Bene_Geo_Lvl	Beneficiary Geography Level	Indicates geography level that the data in the row has been aggregated (e.g. National, State, and County)
Bene_State_Abrvtn	Beneficiary State Abbreviation	State abbreviation of beneficiary residence
Bene_State_Desc	Beneficiary State	Area of beneficiary residence
Bene_County_Desc	Beneficiary County	County of beneficiary residence
Bene_FIPS_Cd	Beneficiary FIPS Code	State/County FIPS code of beneficiary residence
Tot_Benes	Total Beneficiaries	Count of all Medicare beneficiaries
Orgnl_Mdcr_Benes	Original Medicare Beneficiaries	Count of all Original Medicare beneficiaries
MA_and_Oth_Benes	Medicare Advantage and Other Health Plan Beneficiaries	Count of all Medicare Advantage and Other Health Plan Beneficiaries
Aged_Tot_Benes	Total Aged Beneficiaries	Count of Medicare-aged beneficiaries
Aged_ESRD_Benes	Aged ESRD Beneficiaries	Count of Medicare-aged beneficiaries with End Stage Renal Disease

Medicare Enrollment Data		
Aged_No_ESRD_Benes	Aged Beneficiaries Without ESRD	Count of Medicare-aged beneficiaries without End Stage Renal Disease
Dsbld_Tot_Benes	Total Disabled Beneficiaries	Count of Medicare disabled beneficiaries
Dsbld_ESRD_and_ESRD_Only_Benes	Disabled Beneficiaries with ESRD and ESRD Only Beneficiaries	Count of Medicare disabled beneficiaries with End Stage Renal Disease and beneficiaries with End Stage Renal Disease only
Dsbld_No_ESRD_Benes	Disabled Beneficiaries Without ESRD	Count of Medicare disabled beneficiaries without End Stage Renal Disease
A_B_Tot_Benes	Total Medicare Part A and Part B Beneficiaries	Count of Medicare beneficiaries enrolled in Hospital Insurance (Part A) and Supplementary Medical Insurance (Part B)
A_B_Orgnl_Mdcr_Benes	Original Medicare Part A and Part B Beneficiaries	Count of Original Medicare beneficiaries enrolled in Hospital Insurance (Part A) and Supplementary Medical Insurance (Part B)
A_B_MA_and_Oth_Benes	Medicare Advantage and Other Health Plan Part A and Part B Beneficiaries	Count of Medicare Advantage and Other Health Plan Beneficiaries enrolled in Hospital Insurance (Part A) and Supplementary Medical Insurance (Part B)
Prscrptn_Drug_Tot_Benes	Total Medicare Part D beneficiaries	Count of all Medicare Prescription Drug (Part D) beneficiaries
Prscrptn_Drug_PDP_Benes	Total Medicare Prescription Drug Plan beneficiaries	Count of Medicare Prescription Drug (Part D) beneficiaries enrolled in a Prescription Drug Plan
Prscrptn_Drug_MAPD_Benes	Total Medicare Advantage Prescription Drug Plan	Count of Medicare Prescription Drug (Part D) beneficiaries enrolled in a Medicare Advantage Prescription Drug

Medicare Enrollment Data		
	beneficiaries	Plan
Prscrptn_Drug_Deemed_Eligible_Full _LIS_Benes	Total Part D Deemed Eligible Full LIS Beneficiaries	Count of certain groups of Medicare beneficiaries who are automatically deemed eligible for the low-income subsidy (LIS) and do not have to apply.
Prscrptn_Drug_Full_LIS_Benes	Total Part D Full LIS Beneficiaries	Count of Medicare beneficiaries with limited income and resources who do not fall into one of the deemed subsidy groups and are enrolled for full subsidy. These beneficiaries successfully applied for a low-income subsidy (LIS).
Prscrptn_Drug_Partial_LIS_Benes	Total Part D Partial LIS Beneficiaries	Count of Medicare beneficiaries with limited income and resources who do not fall into one of the deemed subsidy groups and are enrolled for partial subsidy. These beneficiaries successfully applied for a low-income subsidy (LIS).
Prscrptn_Drug_No_LIS_Benes	Total Part D Beneficiaries with No LIS	Count of Medicare Part D beneficiaries with no low-income subsidy (LIS).

Medicare Drugs Data		
Variable Name	**Term description**	**Definition**
Brnd_Name	Brand Name	The name of the brand for the drug.
Gnrc_Name	Generic Name	The generic name of the drug.

Medicare Drugs Data		
Tot_Mftr	Total Manufacturers	Number of manufacturers for the drug.
Mftr_Name	Manufacturer Name	Name of the manufacturer.
Tot_Spndng_(year)	Total Spending for (year)	Total spending on the drug in that particular year.
Tot_Dsg_Unts_(year)	Total Dosage Units for (year)	Total dosage units of the drug dispensed in that particular year.
Tot_Clms_(year)	Total Claims for (year)	Total number of claims filed for the drug in that particular year.
Tot_Benes_(year)	Total Beneficiaries for (year)	Total number of beneficiaries who received the drug in that particular year.
Avg_Spnd_Per_Dsg_Unt_Wghtd_(year)	Avg Spend Per Dosage Unit Weighted for (year)	Average spending per dosage unit, weighted, in that particular year.
Avg_Spnd_Per_Clm_(year)	Average Spend Per Claim for (year)	Average spending per claim for the drug in that particular year.
Outlier_Flag_(year)	Outlier Flag for (year)	Indicator if the drug spending or usage was an outlier in that particular year.
Chg_Avg_Spnd_Per_Dsg_Unt_20_21	Change in Average Spend Per Dsg Unit 20-21	Change in average spending per dosage unit from 2020 to 2021.
CAGR_Avg_Spnd_Per_Dsg_Unt_17_21	CAGR Average Spend Per Dsg Unit 17-21	Compound annual growth rate of the average spending per dosage unit from 2017 to 2021.

Medicare Physicians Data		
Variable Name	**Term Description**	**Definition**
Rfrg_NPI	Referring NPI	National Provider Identifier (NPI) for the referring physician.
Rfrg_Prvdr_Last_Name_Org	Referring Provider Last Name/ Organization	Last name or organization name of the referring provider.
Rfrg_Prvdr_First_Name	Referring Provider First Name	First name of the referring provider.
Rfrg_Prvdr_MI	Referring Provider Middle Initial	The middle initial of the referring provider.
Rfrg_Prvdr_Crdntls	Referring Provider Credentials	Credentials of the referring provider (e.g., M.D., D.O.)
Rfrg_Prvdr_Gndr	Referring Provider Gender	Gender of the referring provider.
Rfrg_Prvdr_Ent_Cd	Referring Provider Entity Code	Entity code of the provider (e.g., individual "I" or organization “O”).
Rfrg_Prvdr_St1	Referring to Provider Street 1	The first line of the street address for the referring provider.
Rfrg_Prvdr_St2	Referring to Provider Street 2	The second line of the street address for the referring provider.
Rfrg_Prvdr_City	Referring to Provider City	City of the referring provider's practice location.
Rfrg_Prvdr_State_Abrvtn	Referring to Provider State Abbreviation	State the abbreviation where the referring provider is located.
Rfrg_Prvdr_State_FIPS	Referring to Provider State FIPS	Federal Information Processing Standards code for the state.
Rfrg_Prvdr_Zip5	Referring to Provider ZIP5	First 5 digits of the ZIP code for the referring provider's location.
Rfrg_Prvdr_RUCA_CAT	Referring to Provider RUCA Category	Rural-Urban Commuting Area (RUCA) category for the location.
Rfrg_Prvdr_RUCA	Referring to Provider RUCA Code	RUCA code for the referring provider's location.
Rfrg_Prvdr_RUCA_Desc	Referring to Provider RUCA Description	Description of the RUCA location for the provider.

Rfrg_Prvdr_Cntry	Referring to Provider Country	Country where the referring provider is located.
Rfrg_Prvdr_Type_cd	Referring to Provider Type Code	Code for the type of provider.
Rfrg_Prvdr_Type	Referring to Provider Type	Type of provider (e.g., Internal Medicine, Family Practice).
Rfrg_Prvdr_Type_Flag	Referring Provider Type Flag	Flag indicating specifics about the provider type (e.g., S for specialist).
BETOS_Lvl	BETOS Level	Level of the Berenson-Eggers Type of Service (BETOS) coding.
BETOS_Cd	BETOS Code	Berenson-Eggers Type of Service code.
BETOS_Desc	BETOS Description	Description of the BETOS code.
HCPCS_CD	HCPCS Code	Healthcare Common Procedure Coding System code.
HCPCS_Desc	HCPCS Description	Description of the HCPCS code.
Suplr_Rentl_Ind	Supplier Rental Indicator	Indicator if the item is a rental (Y for yes, N for no).
Tot_Suplrs	Total Suppliers	Total number of suppliers providing the service or item.
Tot_Suplr_Benes	Total Supplier Beneficiaries	Total number of beneficiaries receiving the service or item from suppliers.
Tot_Suplr_Clms	Total Supplier Claims	Total number of claims submitted for the service or item.
Tot_Suplr_Srvcs	Total Supplier Services	Total number of services provided.
Avg_Suplr_Sbmtd_Chrg	Average Supplier Submitted Charge	Average charge submitted by suppliers for the service or item.
Avg_Suplr_Mdcr_Alowd_Amt	Average Supplier Medicare Allowed Amount	Average amount Medicare allows for the service or item.
Avg_Suplr_Mdcr_Pymt_Amt	Average Supplier Medicare Payment Amount	Average amount Medicare pays to the supplier for the service or item.
Avg_Suplr_Mdcr_Stdzd_Amt	Average Supplier Medicare Standardized Amount	Average standardized amount for Medicare payments adjusted to remove geographic differences in payment rates across different parts of the country.

Technology used:

Python and Excel

Data Preprocessing: Python was utilized to perform the data cleaning and transformation of large datasets, ensuring high-quality, consistent data. Excel was used for reviewing data, performing simpler transformations, and quickly verifying results.

Tableau

Visualization and Analysis: Tableau was used to create dynamic visualizations that facilitated a deeper understanding of the data patterns and trends. These visualizations were instrumental in the analysis phase, allowing us to interact with the data and derive actionable insights.

Data Pre-processing:

Our primary objectives in preprocessing were to ensure data quality, facilitate easier data analysis, and prepare the dataset for visualization in Tableau.

1. Data Preprocessing for Medicare Dataset (2013 to 2023):

The initial stage of our data preprocessing involved loading the Medicare dataset from 2013 to 2023, sourced from data.gov - **medicare_2013_to_2023.csv**.

Steps and Techniques:

I. **Excluding incomplete data:** For the year 2023 data was not available for the months nearing the end so all data for the year 2023 was removed.
II. **Removing Unnecessary Columns:** Eliminated redundant columns like 'bene_geo_lvl' which had consistent values ("County"), providing no additional insight.
III. **Data consistency:** Accumulated columns such as data for the entire ‘Country’ or ‘State’ or ‘Year’ were removed to finalize data in a monthly and county-wise format for correct means of aggregation during visualization.
IV. **Dropping unknowns values:** Several states having ‘Unknown’ counties were dropped.

V. **Adding features:** Developed new features such as 'region' from 'state' using a mapping function for regional trends in Medicare usage. However, this was not used in the final analysis as the trends did not vary significantly among regions.

VI. **Renaming columns:** Several columns were renamed for quick understanding

Output: medicare_final.csv

2. Data Pre-processing for Drugs dataset:

Steps and Techniques:

1. **Dropping aggregated columns:** Aggregated rows such as 'Overall' manufacturer information for each drug were removed.
2. **Data transformation:** The dataset contains many columns for each year such as avg_spend_2017, and avg_spend_2018 for all the years from 2017 to 2022. This was transformed in such a way that only one column exists for one variable (avg_spend) and the data for each year is captured with year values in a single column.

Output: drugdata_transformed.csv

3. Data Preprocessing for Physicians Dataset:

The **'physicians_data.csv'** dataset contains detailed information about healthcare providers in various locations and the types of medical equipment prescribed by them. **'physicians_data_now.csv'** is obtained after Excel's default preprocessing for handling leading zeros etc,

Steps and Techniques:

I. **Filtering Data by Country:** Dataset was filtered to include only records where the referring provider's country was 'US', focusing our analysis on US-based physicians.

II. **Cleaning Credential Information:** The credentials data was cleaned by removing any periods and standardizing the format to prevent data duplication.

III. **Renaming and Dropping Columns:** Unnecessary columns such as street names, columns with numerous subcategories such as HCPCS (a subcategory of BETOS) were removed, and the remaining columns were renamed.

IV. **Handling Missing Values:** Columns with substantial missing values such as 'Tot_Suplr_Benes' for which beneficiary count below 11 was not entered to protect beneficiary privacy were dropped.

Output: n_physicians_updated.csv (has multiple occurrences of physicians since the dataset is specific to BETOS equipment prescribed by them)

2. 1 Continued Data Preprocessing for Physicians Dataset:

Further cleaning and transformations were applied to n_physicians_updated.csv to make it physician-specific instead of BETOS-specific.

Steps and Techniques:

I. **Filter data:** Columns specific to physicians were filtered and the remaining columns specific to BETOS medical equipment were dropped.

II. **Aggregating Data by each physician:** The data is grouped by the National Provider Identifier (NPI), and the first occurrence of each NPI is taken to remove any potential duplicates. This step ensures that each physician's data is unique.

Output: n_physicians_only.csv (specific to physicians only)

2. 2 Processing Physicians Data for Female Physicians:

Steps and Techniques:

I. **Aggregate over each state:** The dataset specific to physicians was aggregated upon each 'State' to have data specific to female physicians such as the number of female and total number of physicians which is used to calculate the percentage of female physicians in each state in the end.

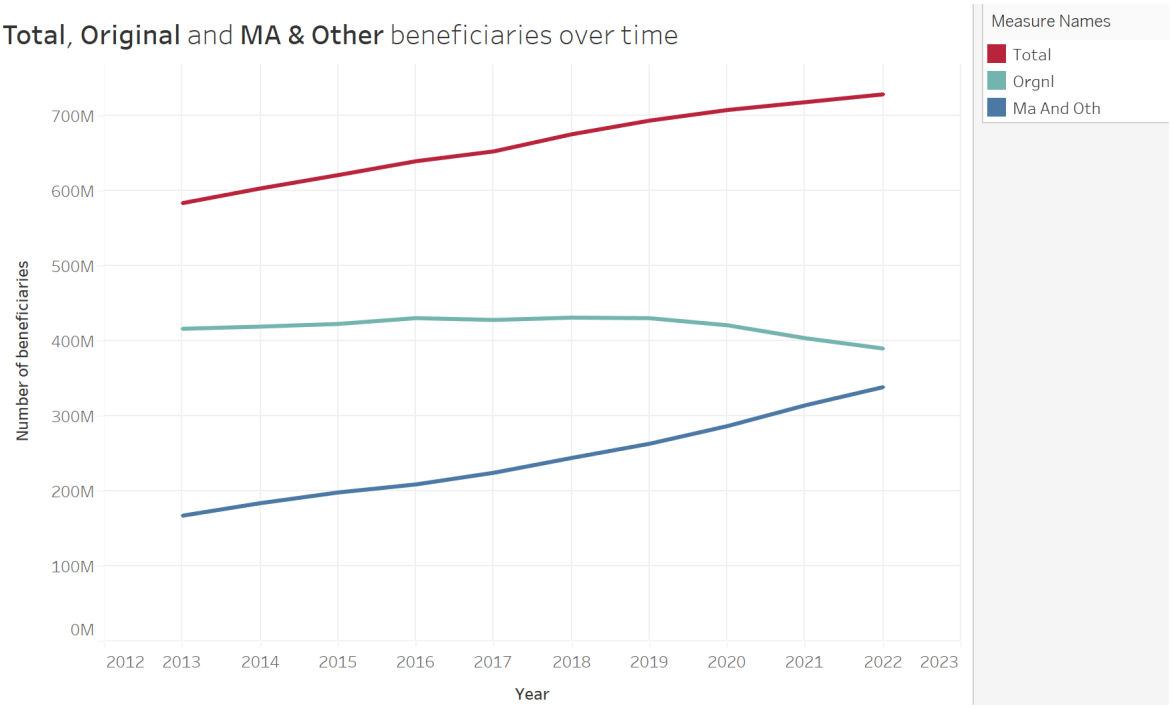
Output: n_female_percentage.csv

Data Analysis:

1. Medicare Dataset Analysis:

Visualization 1: Total, Original, and MA & Other beneficiaries over time

Overview: This line graph visualizes trends in Medicare beneficiary enrollment. Total enrollment, Original Medicare, and Medicare Advantage (MA) & Other plans from 2012 to 2022.

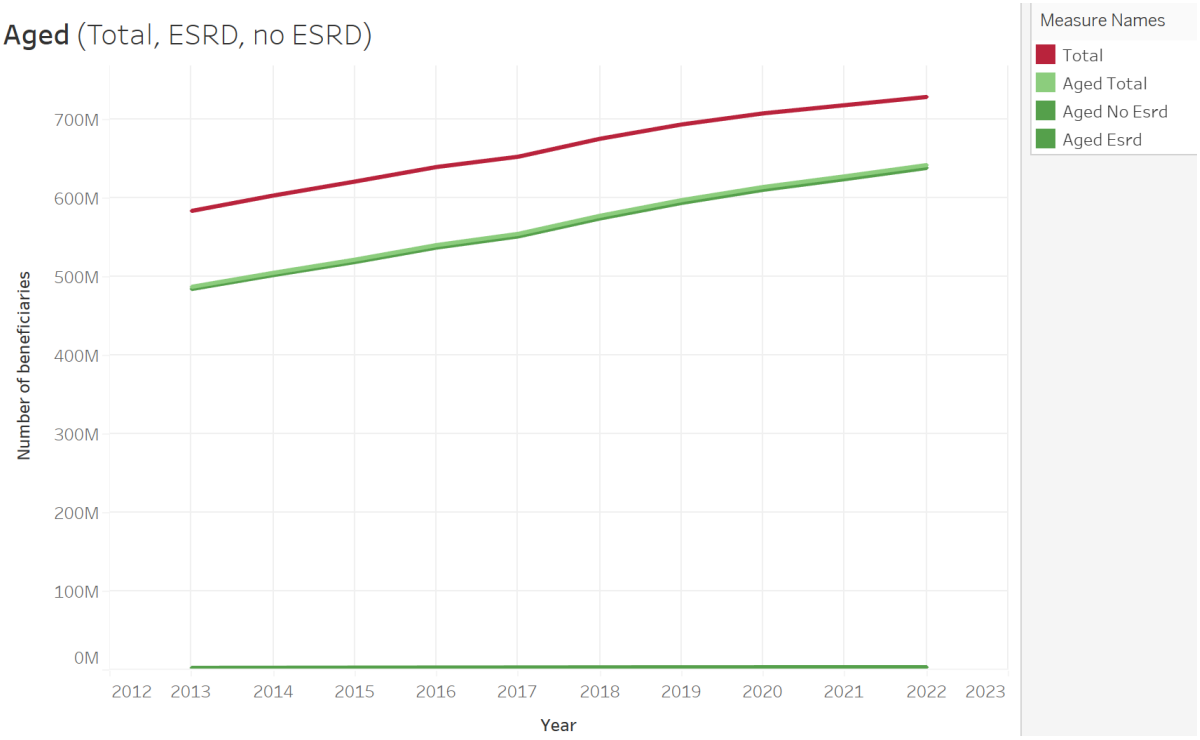


Statistical Analysis:

- The data shows a consistent upward trend in total enrollment over time, with MA and Other plans growing at a faster rate than Original Medicare. This suggests a shifting preference towards these plans.
- The slope of the MA & Other plans line is significantly steeper, indicating a higher rate of change year-over-year.

Visualization 2: Aged (Total, ESRD, no ESRD)

Overview: This line chart disaggregates the aged beneficiary population into those with and without End-Stage Renal Disease (ESRD), over the same timeframe.

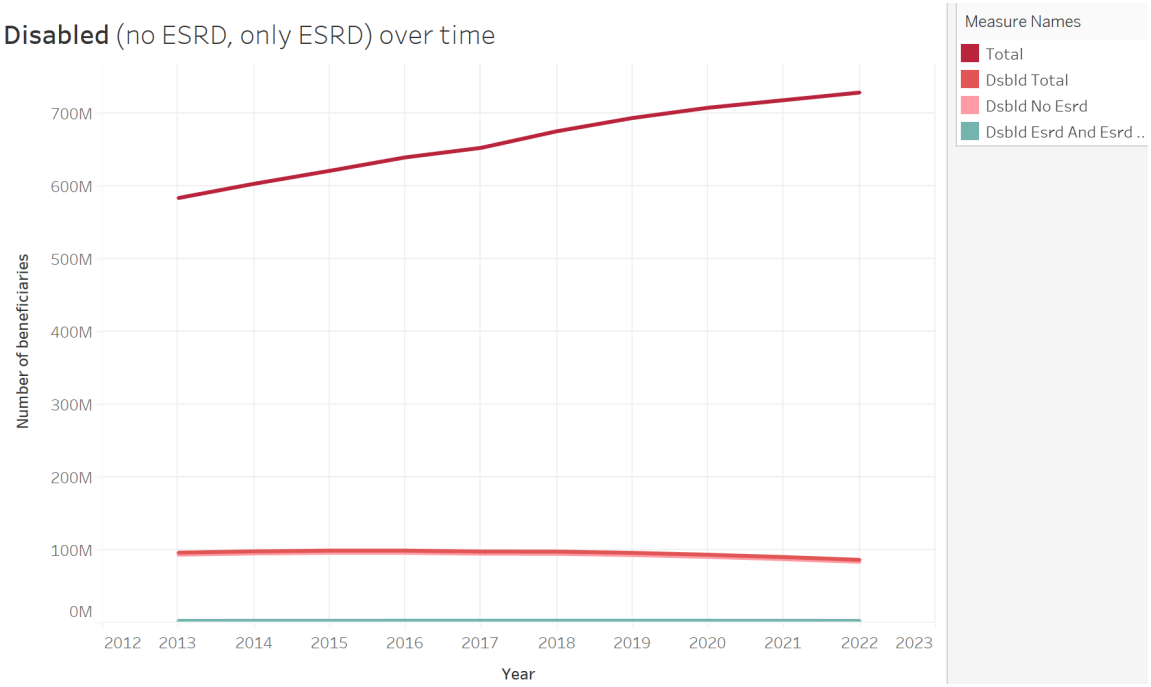


Statistical Analysis:

- Both the total aged beneficiaries and those without ESRD show a consistent increase from 2012 to 2022, indicating an expanding elderly Medicare population.
- The close alignment of the aged total and non-ESRD lines suggests that most of the growth in aged beneficiaries is driven by those without ESRD.

Visualization 3: Disabled (no ESRD, only ESRD) over time

Overview: Analysis of disabled beneficiaries categorized by the presence of ESRD.



Statistical Analysis:

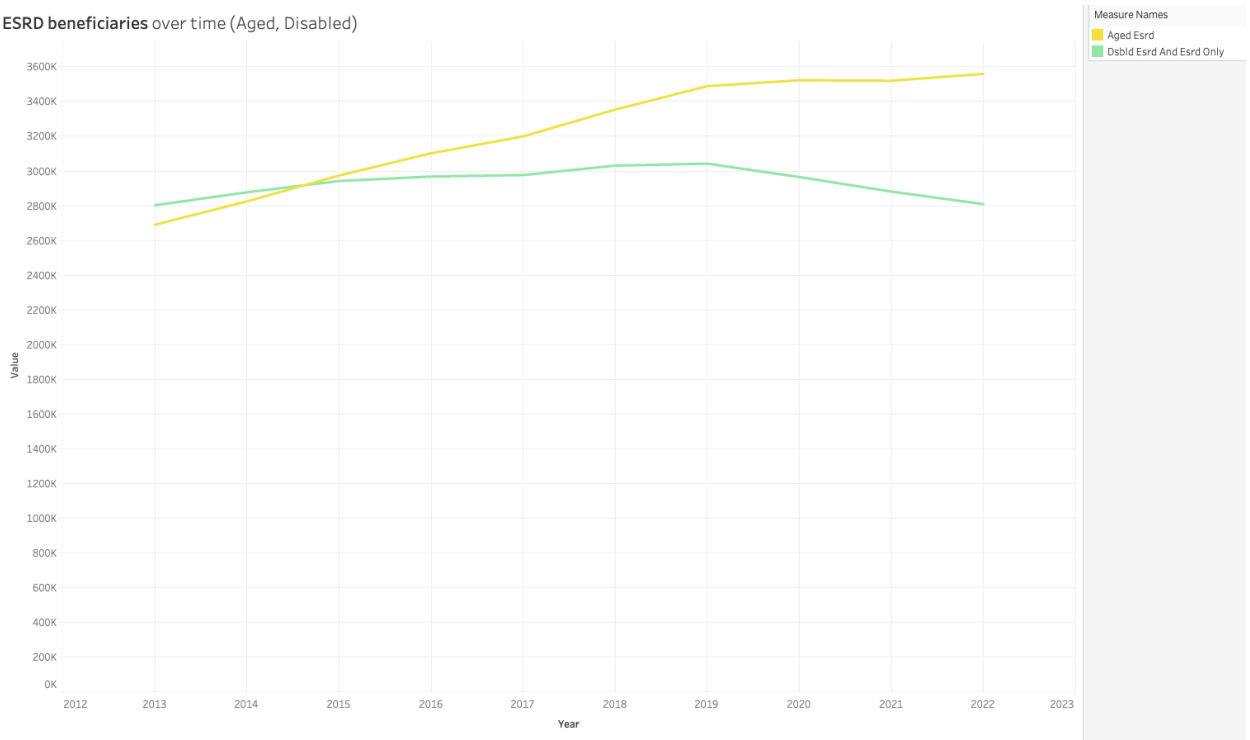
- Disabled Total: This line is very slowly decreasing, indicating very little change over the years, and remains around the 100 million mark.
- Disabled No ESRD: This line is almost merged with the Disabled Total line, suggesting that the majority of disabled beneficiaries do not have End-Stage Renal Disease (ESRD).
- Disabled ESRD and ESRD Only: These two lines are negligible, and barely visible, staying close to zero throughout the observed period.

Cumulative inference:

- From the comparison of the above two trends with the 'Total enrollments' (red line), we can infer that the ratio of Aged to Disabled beneficiaries is almost 4 to 1. In both cases, people with ESRD are very low in number.
- MA and Other plan beneficiaries are growing more and Original plan beneficiaries are growing less over time.
- While the number of aged beneficiaries is growing more over time, the number of disabled beneficiaries is becoming less over time.

Visualization 4: ESRD (Aged, Disabled)

Overview: Trends in ESRD incidence among aged and disabled Medicare beneficiaries.

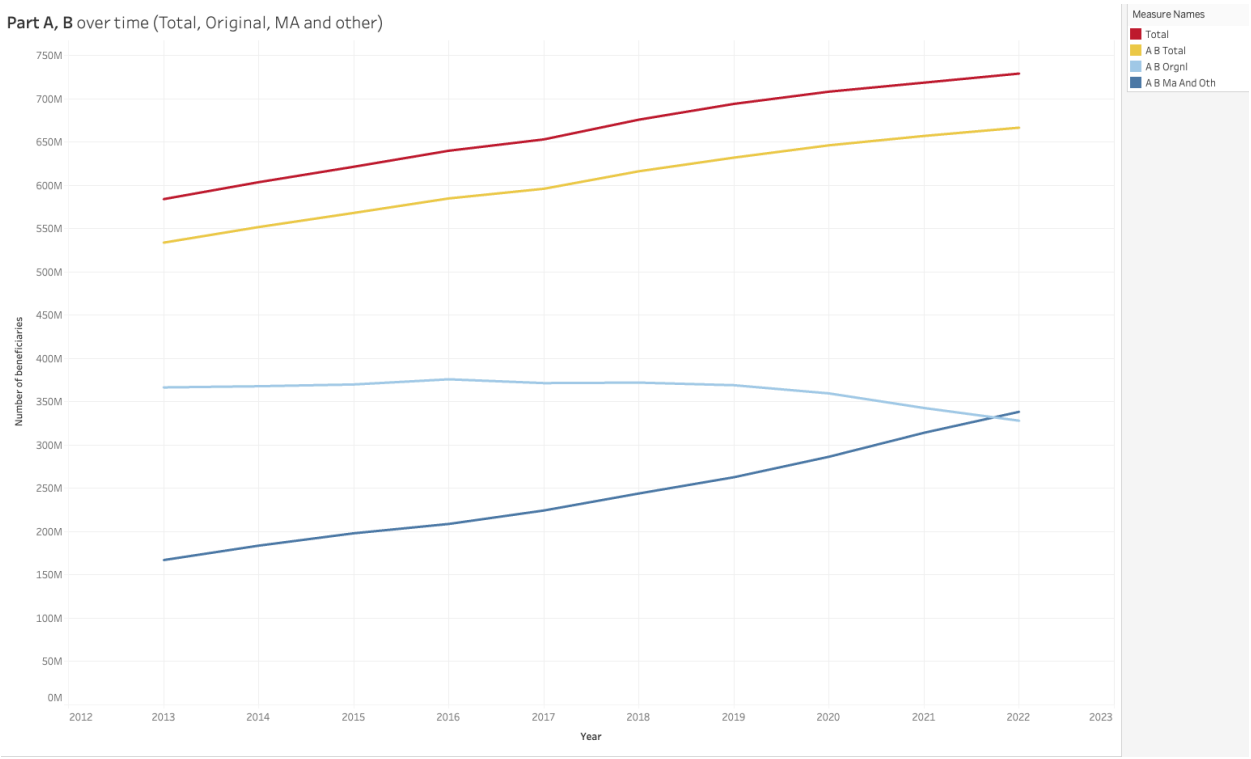


Statistical Analysis:

- A cross-sectional analysis between the two cohorts shows a significant variance in the incidence rates of ESRD. As reflective of our cumulative inference above, the Aged ESRD beneficiaries are growing in number and the disabled ESRD beneficiaries are growing less in number over time

Visualization 5: Part A, B - Hospital Insurance and Supplementary Medical Insurance (total, original, MA and other)

Overview: Enrollment trends in Medicare Parts A and B, comparing Original Medicare and Medicare Advantage and Others.

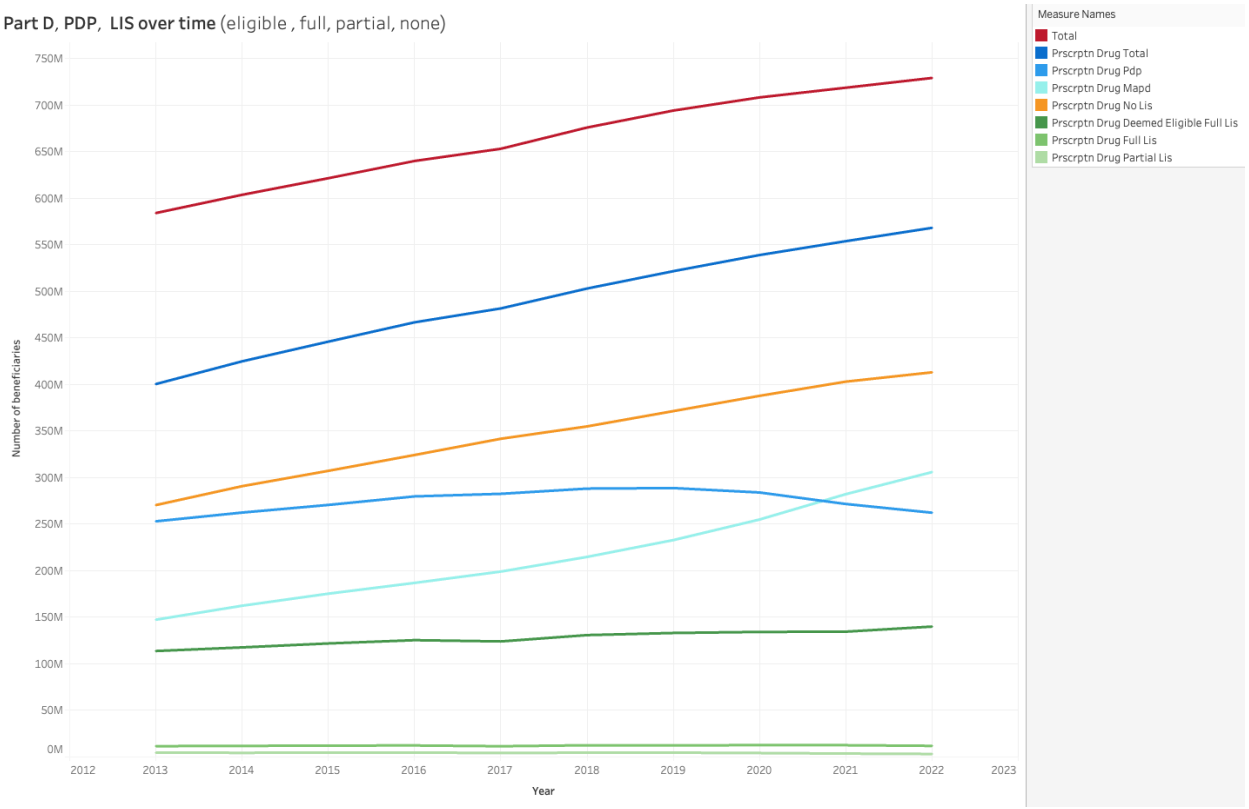


Statistical Analysis:

- The trends for Total and Part A B enrollments exhibit smooth, and consistent growth over time. As expected, the MA and Other Plan is growing to be more popular than the Original Plan.

Visualization 6: Part D (Prescription Drug) and its LIS (Low-Income Subsidy)

Overview: Enrollment dynamics in Medicare Part D, focusing on Prescription Drug Plans (PDP) and Low-Income Subsidy (LIS) eligibility statuses.



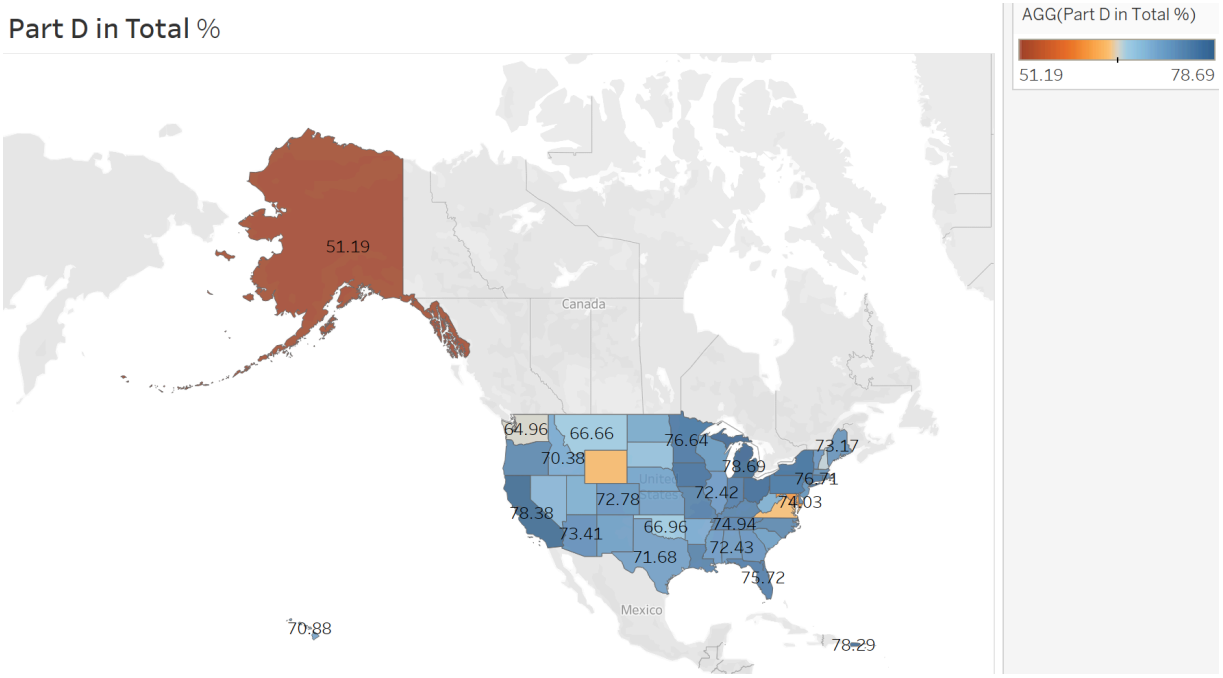
Statistical Analysis:

- Part D beneficiaries (indicated by the first dark blue line) are almost two-thirds of the Total enrollments (red line) and it consistently increases alongside it which implies that most new beneficiaries opt for Part D when they enroll.
- Part D original and MA beneficiaries (indicated by the other two blue lines below) reflect the preference of beneficiaries for MA and Other Plan over the Original Plan as we have seen in all the other charts above.
- No LIS (orange line) beneficiaries are those who do not qualify for any form of Low Income Subsidy. This is almost half of the Total beneficiaries throughout the years which indicates that almost half of Medicare beneficiaries might not need additional financial assistance.
- Among the other types of LIS: Deemed Eligible Full, Full, and Partial (the three green lines in order). Almost one-fourth of the beneficiaries automatically qualify for LIS (Deemed Eligible Full - first green line) and very few people qualify for Full and Partial LIS.

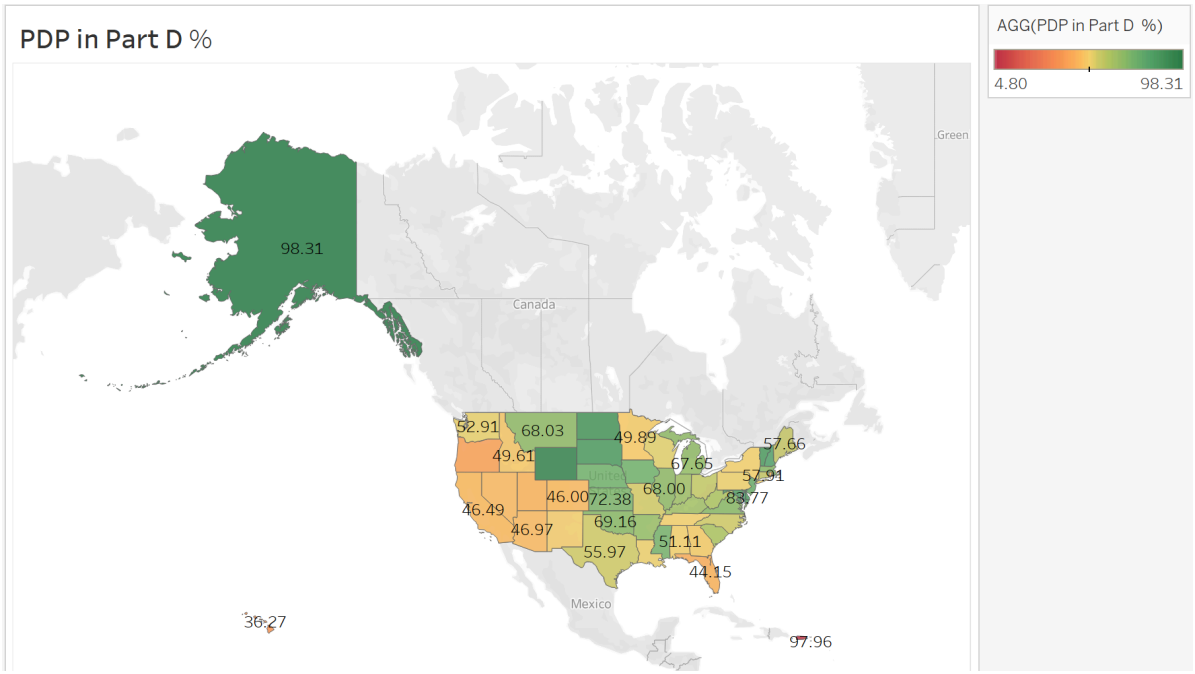
Geographical Analysis of Medicare Part D Coverage across the US:

Overview: The series of choropleth maps visualize several dimensions of Medicare Part D coverage: total enrollment percentage, Prescription Drug Plans (PDP) enrollment, Medicare Advantage Prescription Drug (MAPD) plan enrollment, and the percentage of beneficiaries not receiving the Low-Income Subsidy (LIS).

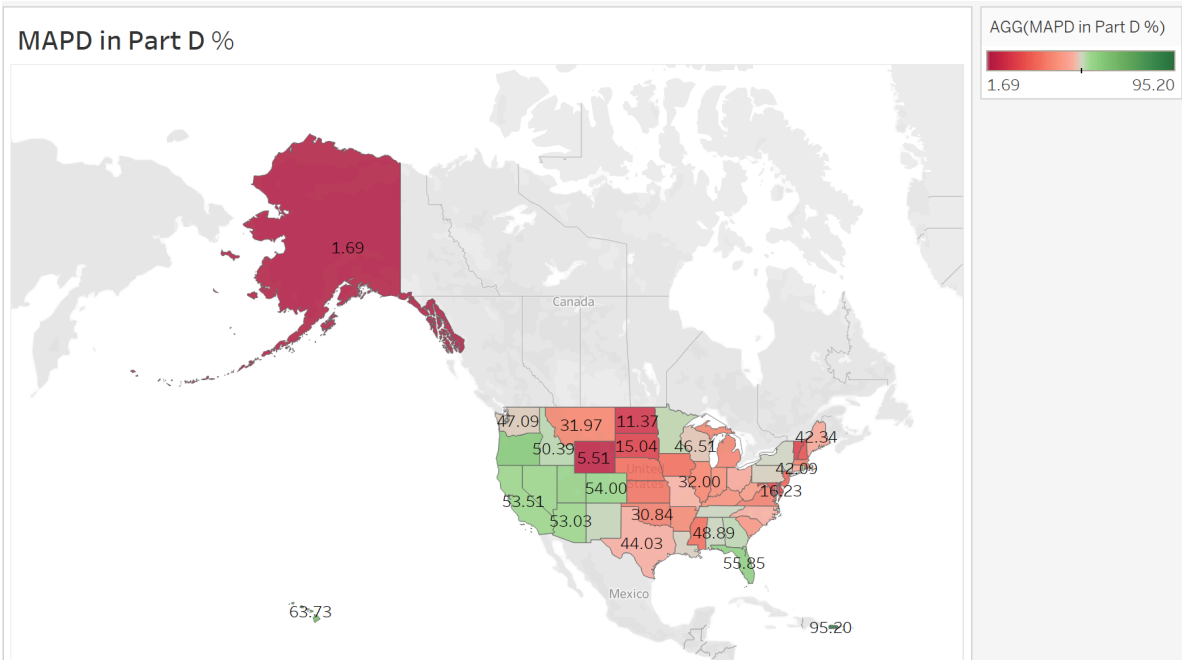
Visualization 7: Part D enrollment across the US:



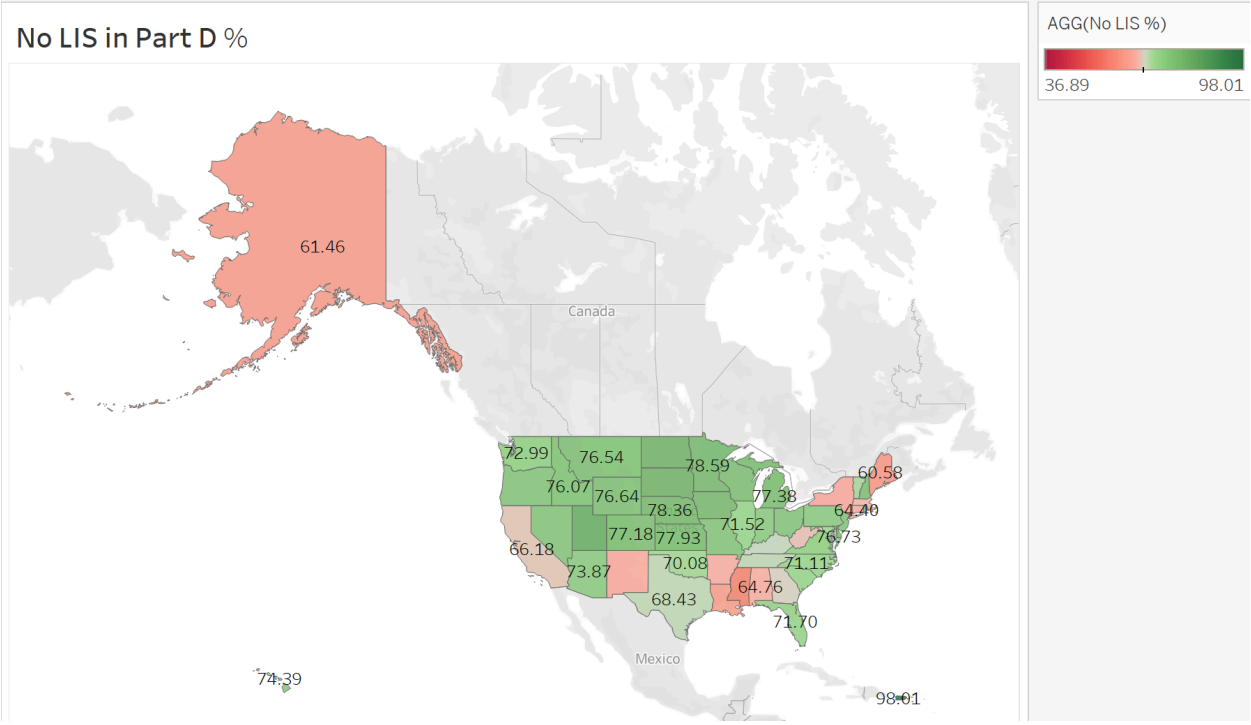
- Apart from a half full of states almost all the states have over 65% enrollment in part D. Alaska has the lease enrollment which is about 50%.



- Even though the Part D enrollment was only about 50% in Alaska, almost everyone in Part D in Alaska is enrolled under the PDP Plan which constitutes the Original Part D Plan. The central America has a higher enrollment in the PDP Plan.



- Western America and some other states (marked in green) Part D enrollments are highly associated with MAPD which is the MA and Other Plan of Part D.



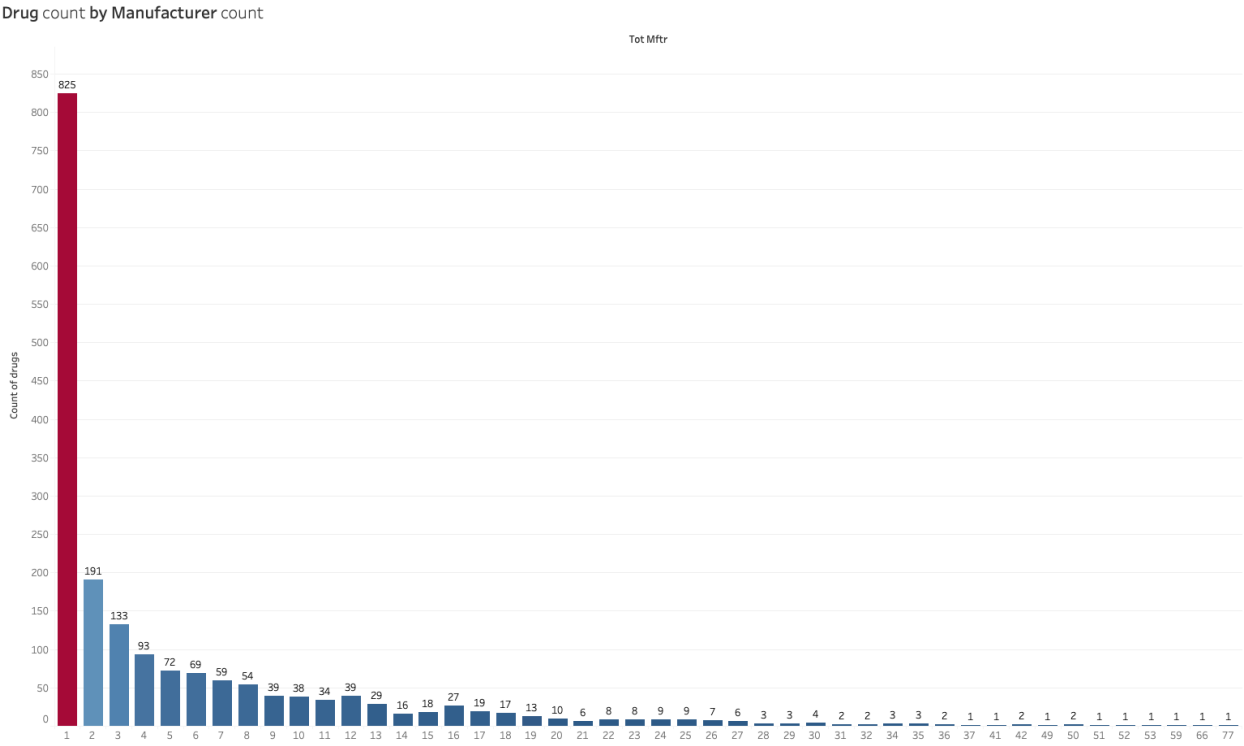
- Except for Alaska and a few southern and eastern states, almost none of them have a higher eligibility for LIS. This could indicate that the majority of the people in these states are generally better off than the others so they do not qualify for LIS.

2. Drugs Dataset Analysis:

The series of visualizations from the drugs dataset provides a detailed overview of drug counts by manufacturer, spending by top brands, and distribution trends over time.

Visualization 8: Drug Count by Manufacturer Count:

Overview: This histogram displays the count of drugs associated with the total number of manufacturers available for them.

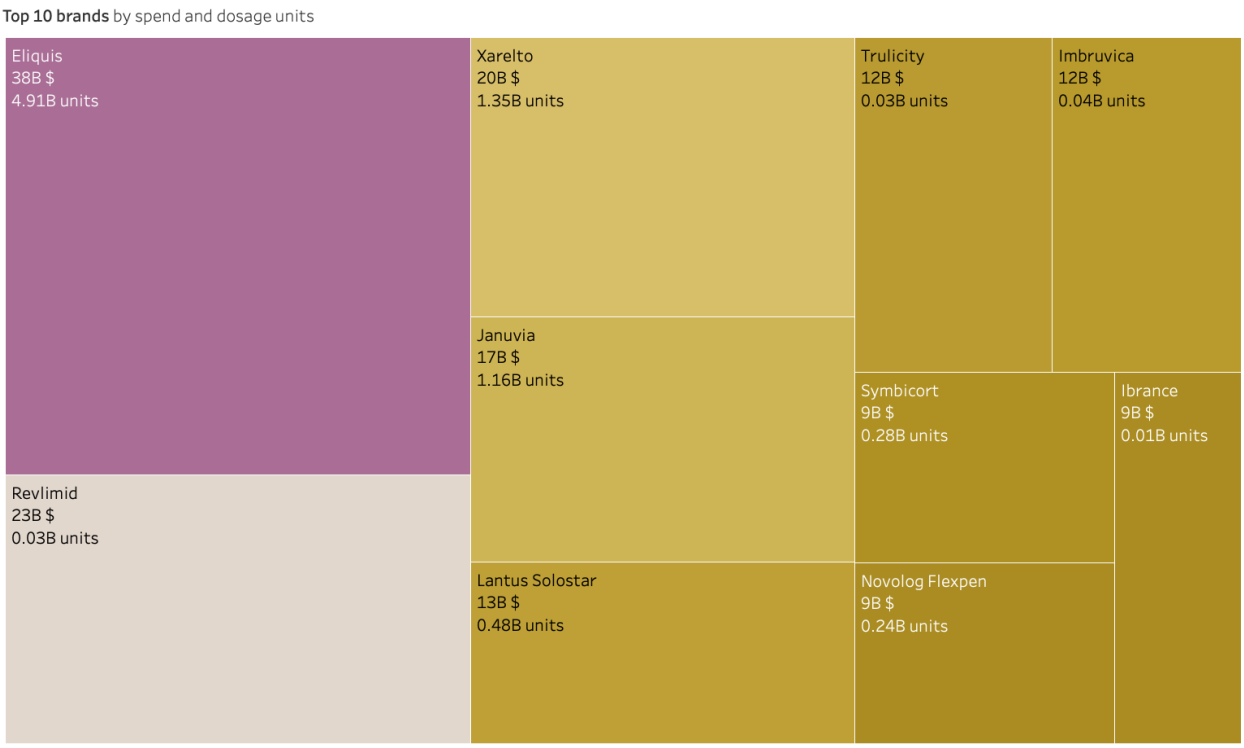


Statistical Analysis:

- The chart predominantly shows that a large number of drugs (over 800) are produced by a single manufacturer, indicating a high level of product specialization and possibly patent control.
- The taller bars on the left end could constitute specialized drugs and the bars on the right end could constitute generic drugs such as Pen Needle which has over 77 manufacturers. This is attributed to its widespread use in treating diabetes, a common condition covered extensively in Medicare.

Visualization 9: Top 10 Brands by Spend and Dosage Units

Overview: This treemap highlights the top 10 drug brands by total spending and dosage units sold.

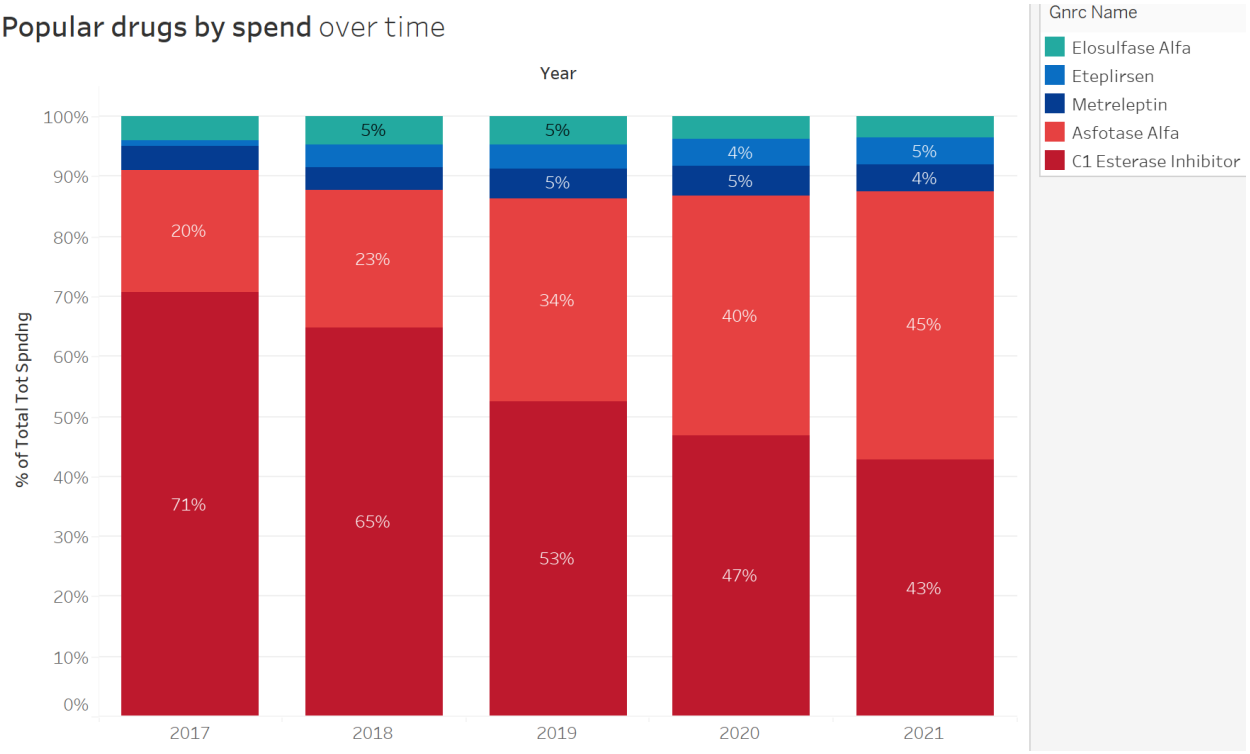


Statistical Analysis:

- The treemap size and color saturation indicate both the financial expenditure and volume of sales. Eliquis leads in both spending and units sold, with a total spend of $38 billion and sales of 4.91 billion units.
- Eliquis, the top brand, shows about 4 times the spending and nearly 500% of the units sold compared to Ibrance, the brand with the 10th highest spending.
- There is no information regarding profit in the dataset, so we cannot conclude that Eliquis despite its highest share of the total spending is the most profitable.

Visualization 10: Popular Drugs by Spend Over Time

Overview: This stacked bar chart shows the proportion of total spending attributed to the top 5 drugs over several years.

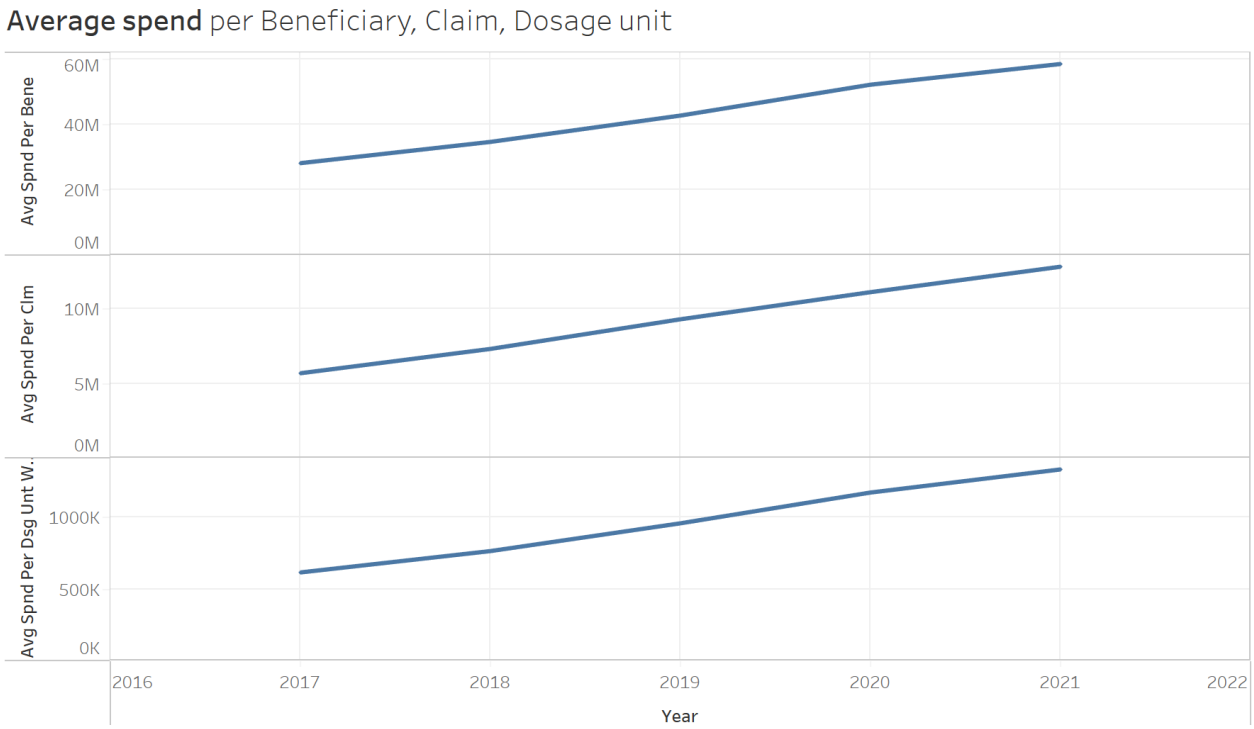


Statistical Analysis:

- The consistent prominence of **C1 Esterase Inhibitor** throughout the years suggests it's a critical component in managing health conditions related to hereditary angioedema (preventing swelling and painful attacks caused by a build-up of fluid resulting from allergic reactions). However, its share has dropped from about 70% to 40% in five years indicating the use of alternative drugs or poor coverage and treatment options associated with the drug leading to the decline in market share.
- The spending on **Asfotase Alfa** indicates growing attention to mineral bone disorders. This trend could reflect a wider recognition of bone health issues among the population or improved coverage and treatment options in health plans, leading to increased usage.

Visualization 11: Average Spend per Beneficiary, Claim, Dosage Unit

Overview: This line graph tracks the average spending per beneficiary, per claim, and per dosage unit over time.

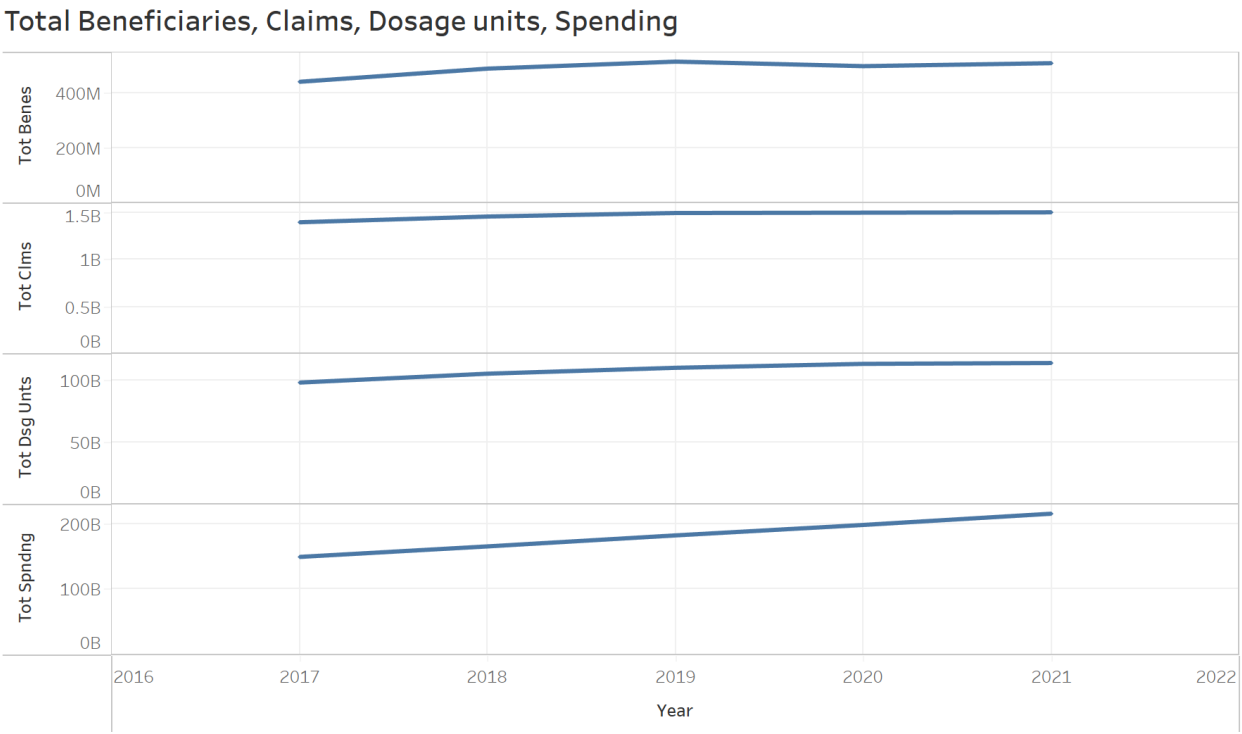


Statistical Analysis:

- The average spending per beneficiary shows a consistent increase from 2016 to 2022. This upward trend suggests a rise in either the cost of treatments per person or more expensive treatments being utilized over the years.
- Similarly, the average spend per claim has also increased steadily over the period. This indicates that the cost per medical claim is rising, which could reflect higher prices for medical services or an increase in the complexity or intensity of treatments per claim.
- The average spend per dosage unit has seen a gradual rise as well, although the slope is less steep compared to the other categories. This suggests a moderate increase in the cost of medication over time.

Visualization 12: Total Beneficiaries, Claims, Dosage Units, Spending

Overview: This set of line charts depicts the total number of beneficiaries, total claims, total dosage units, and total spending over several years.

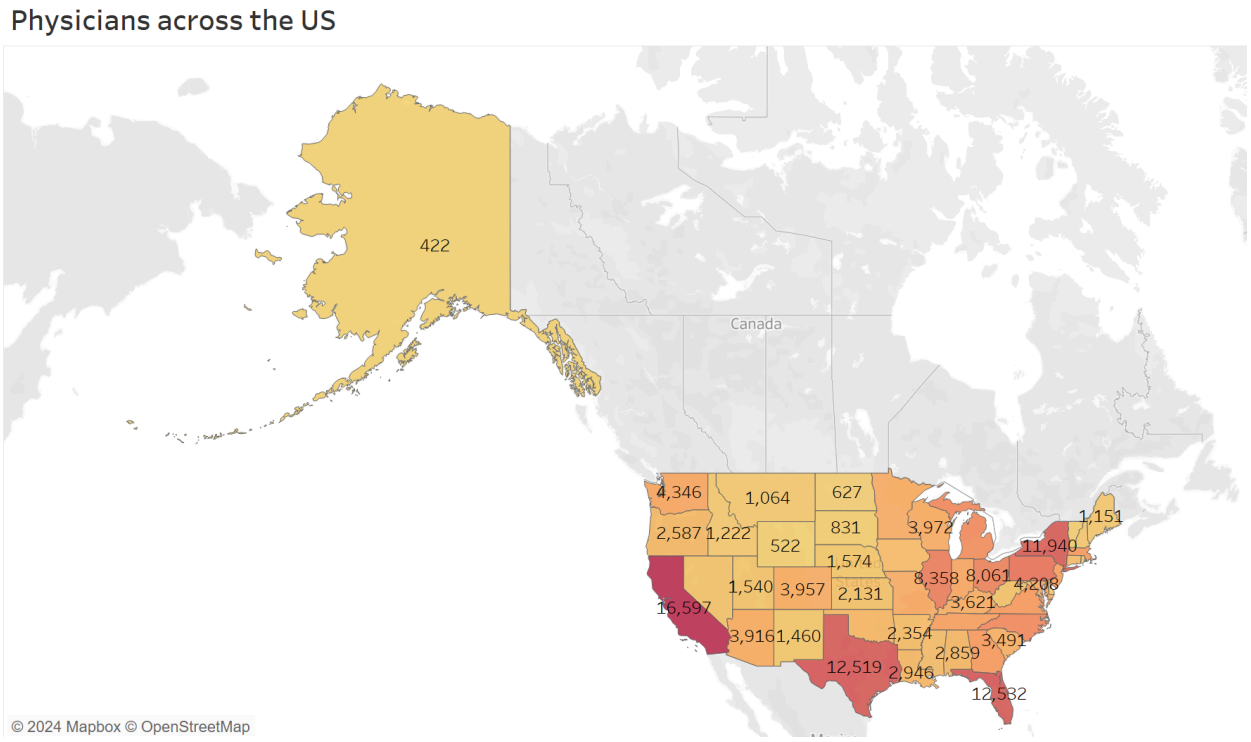


Statistical Analysis:

- The number of beneficiaries and claims has remained relatively stable over the years. The total beneficiaries line appears almost flat, suggesting no significant increase or decrease in the number of individuals covered. There is a noticeable dip around 2019. This could be attributed to the impact of the COVID-19 pandemic.
- Both the total dosage units and total spending show a gradual increase over the period. This upward trend indicates that while the number of beneficiaries and claims has remained stable, the total spending has increased.

3. Physicians and Medical Equipments Dataset Analysis:

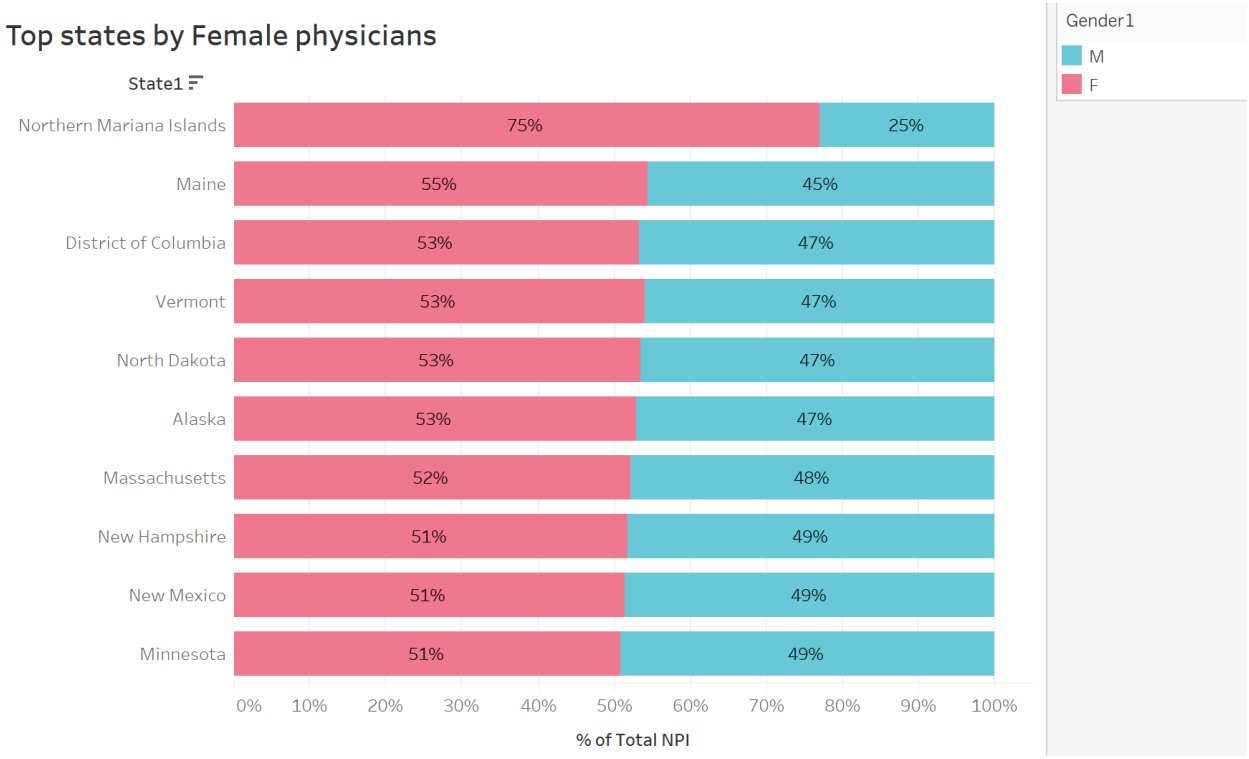
Visualization 13: Distribution of Physicians across the US



Statistical Analysis:

- States like California (16,597), Texas (12,519), and New York (11,940) show significantly higher numbers of physicians compared to other states. This could be due to larger urban populations and advanced healthcare facilities.

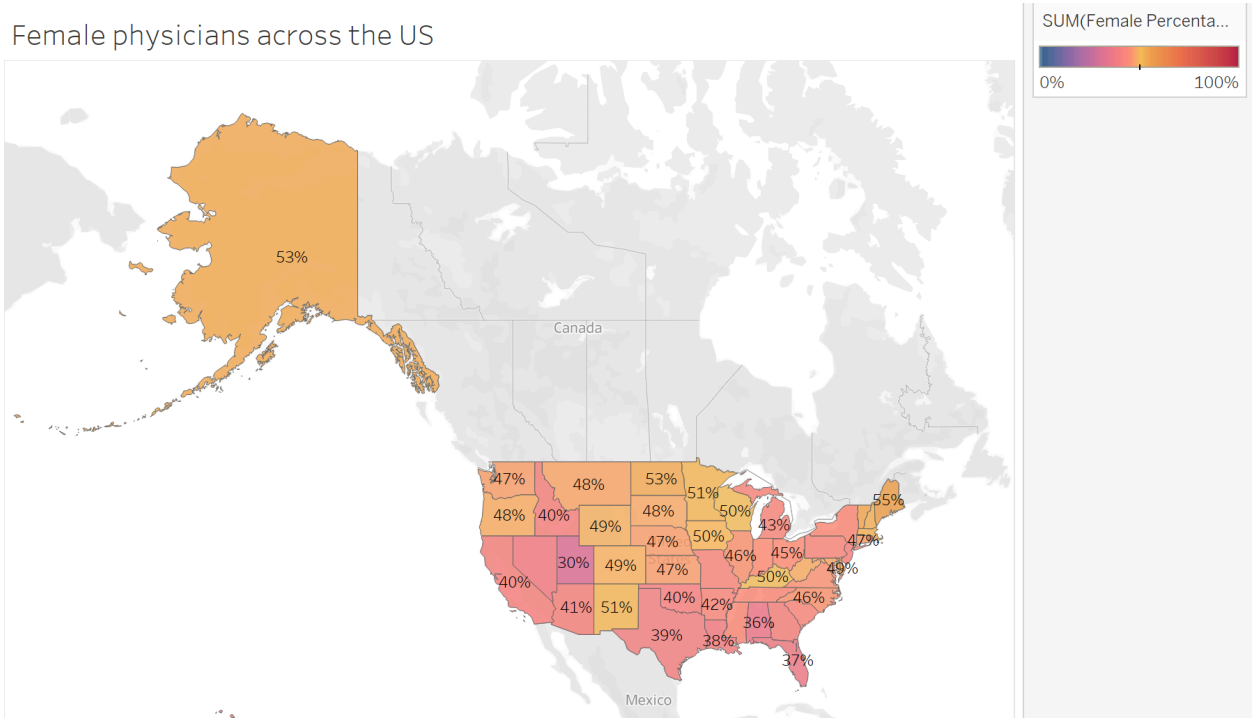
Visualization 14: Top states by Female physicians



Statistical Analysis:

- Northern Mariana Islands: This territory has the highest percentage of female physicians at 75%, which is significantly higher than any other state or territory listed.
- All the top 10 states by female physician distribution have over 50% of females. District of Columbia is in the top 3 states by female physicians percentage.

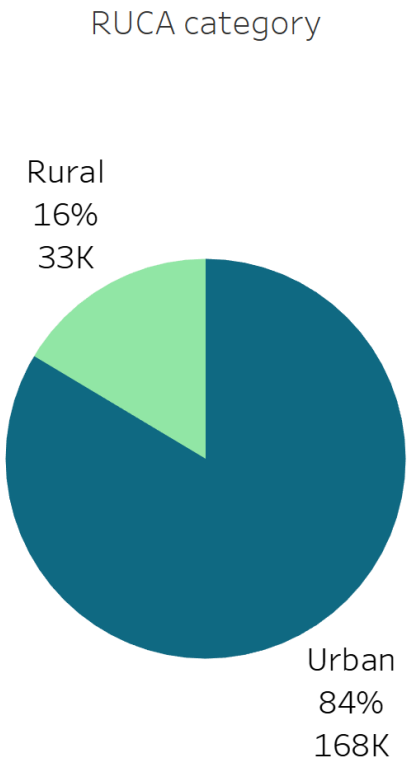
Visualization 15: Distribution of Female Physicians across the US



Statistical Analysis:

- States like Maine (55%) and Alaska (53%) show the highest percentages of female physicians.
- The map shows notable regional differences, with states in the Northeast and some parts of the West (like Alaska) having higher percentages of female physicians. In contrast, several Southern states such as Utah (30%), Mississippi (36%), and Louisiana (37%) have some of the lowest percentages
- States like Colorado, Iowa, and Missouri exhibit an almost equal gender distribution among physicians

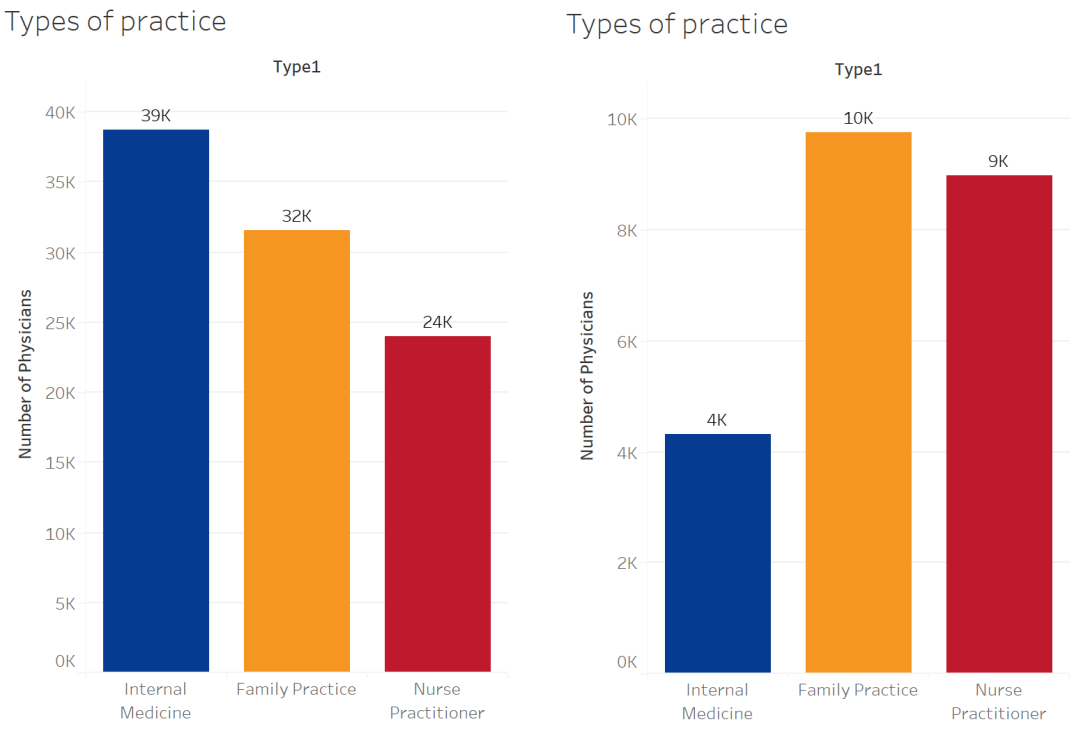
Visualization 16: RUCA Category



Statistical Analysis:

- The ratio of Urban to Rural physicians associated with Medicare is about 85% to 15% which is almost 17:3. This indicates that the number of physicians in the urban sector is over 5 times in comparison to those belonging to the rural sector.

Visualization 17: Distribution of Types of Practice



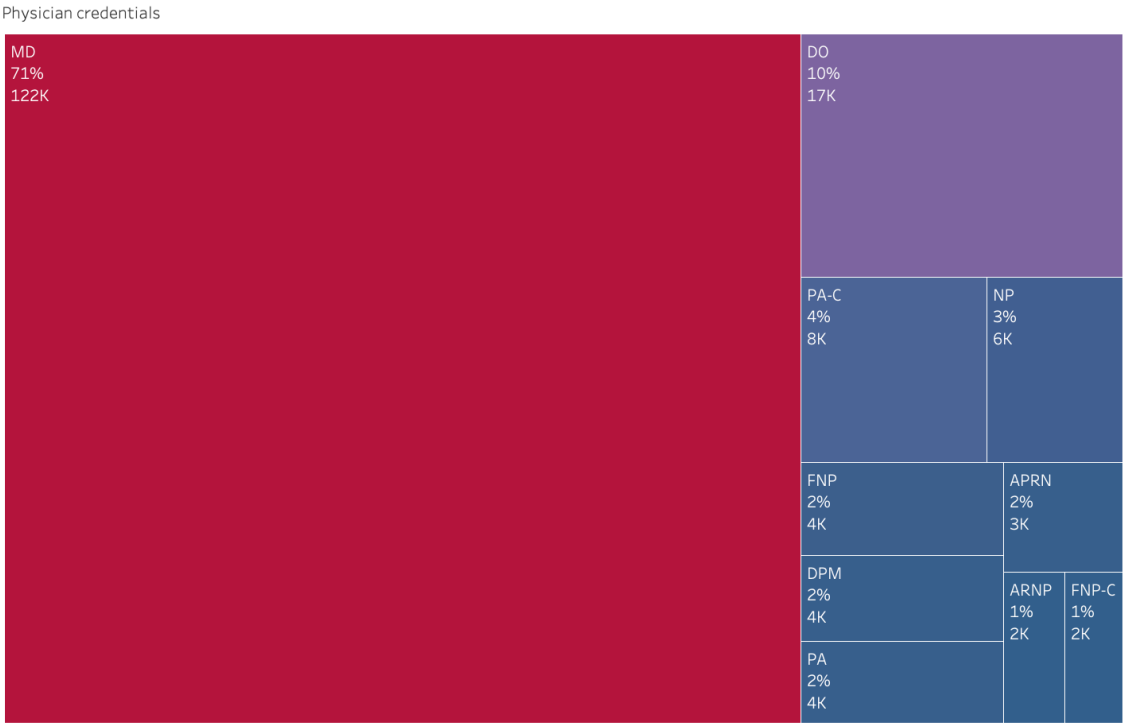
Urban sector

Rural sector

Statistical Analysis:

- Internal medicine, Family practice, and Nurse practitioners are the popular types of practices in both urban and rural sectors.
- However, Internal Medicine is the most popular among the three in urban and least popular in rural areas. Family practice is most popular in rural areas followed by Nurse Practitioners.

Visualization 18: Distribution of Physicians' Credentials



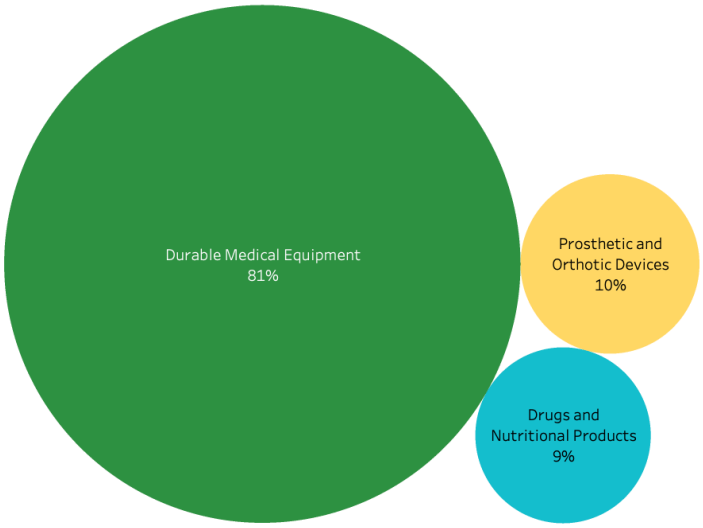
Statistical Analysis:

- MD (Doctor of Medicine): Over 70% of the physicians hold an MD. This high percentage underscores the traditional role of MDs as the primary qualification in the medical profession.
- DO (Doctor of Osteopathic Medicine): Represents 10% of the data with 17,000 individuals.

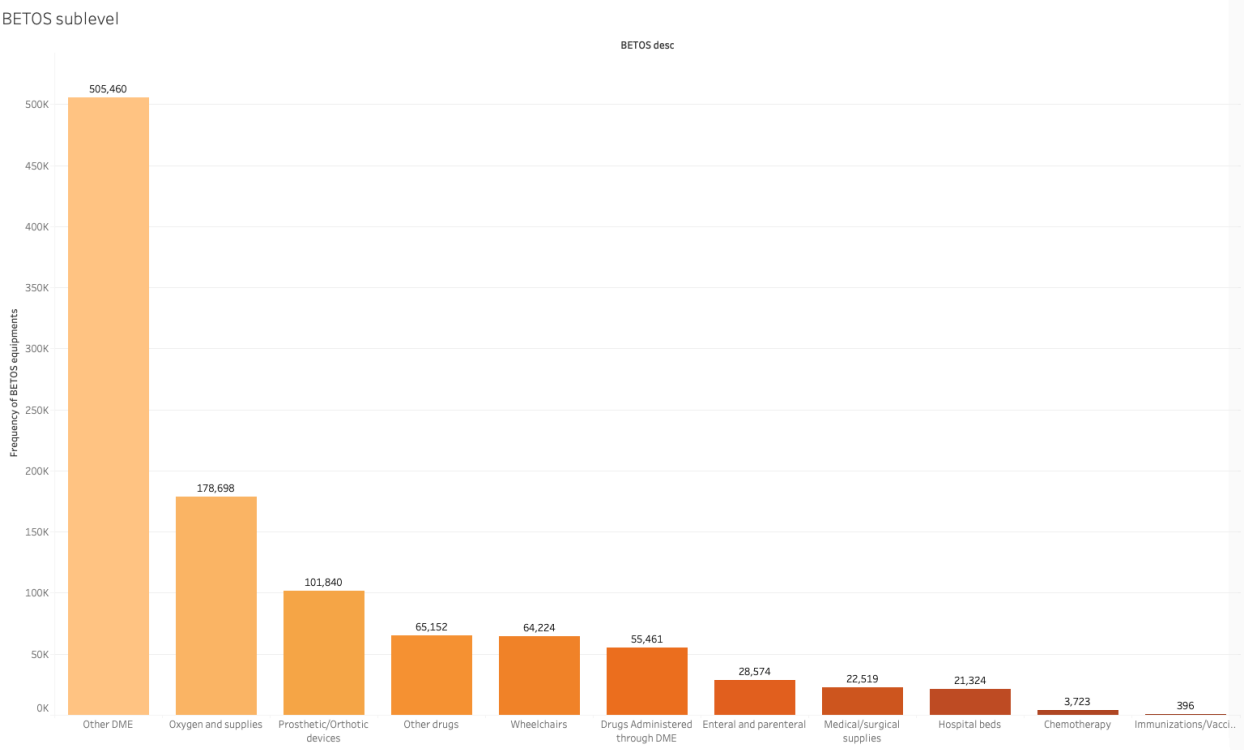
Visualization 19: BETOS (Berenson-Eggers Type of Service) Classification of Medical Product Distribution

Statistical Analysis:

- Durable Medical Equipment: This category constitutes the vast majority (81% or 122K units) of the items in this dataset. It shows that durable medical equipment (like wheelchairs, monitors, etc.) is a predominant need in this healthcare segment.
- Prosthetic and Orthotic Devices, Drugs, and Nutritional Products make up the remaining 10% and 9% of the equipment and products offered by Medicare



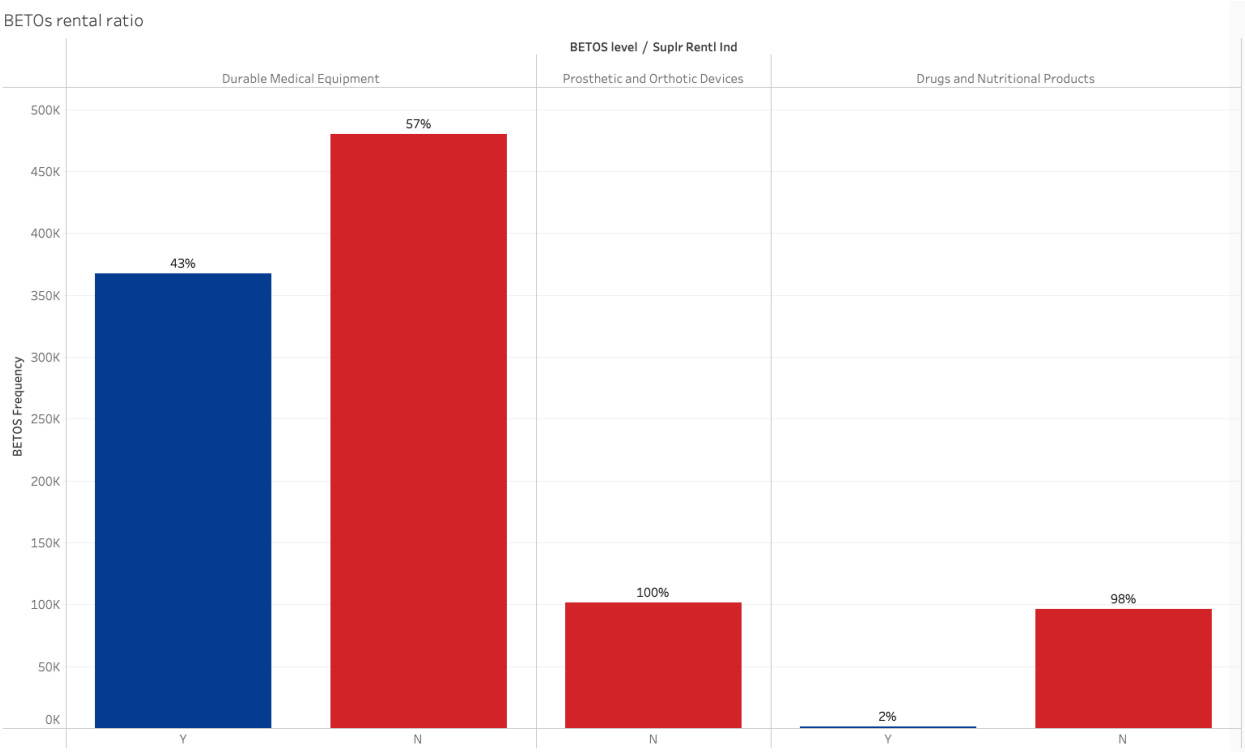
Visualization 20: Distribution of BETOS Medical Products and Services by Category



Statistical Analysis:

- Other DME (Durable Medical Equipment) is the most popular category being prescribed over 500k times. It is almost thrice as frequent as the next category which is Oxygen and Supplies.
- There is a good number of instances of Prosthetics and Wheelchairs with over 100k and 60k prescriptions respectively indicating support and affordability offered related to these services that enable beneficiaries to prefer Medicare in opting for them.
- Specialized services like Chemotherapy and Immunizations/Vaccinations have fewer registrations with just 3k and 300 instances. This could indicate better plans available outside Medicare for such specialized medical conditions.

Visualization 21: Rental and Non-Rental Distribution of BETOS Medical Products



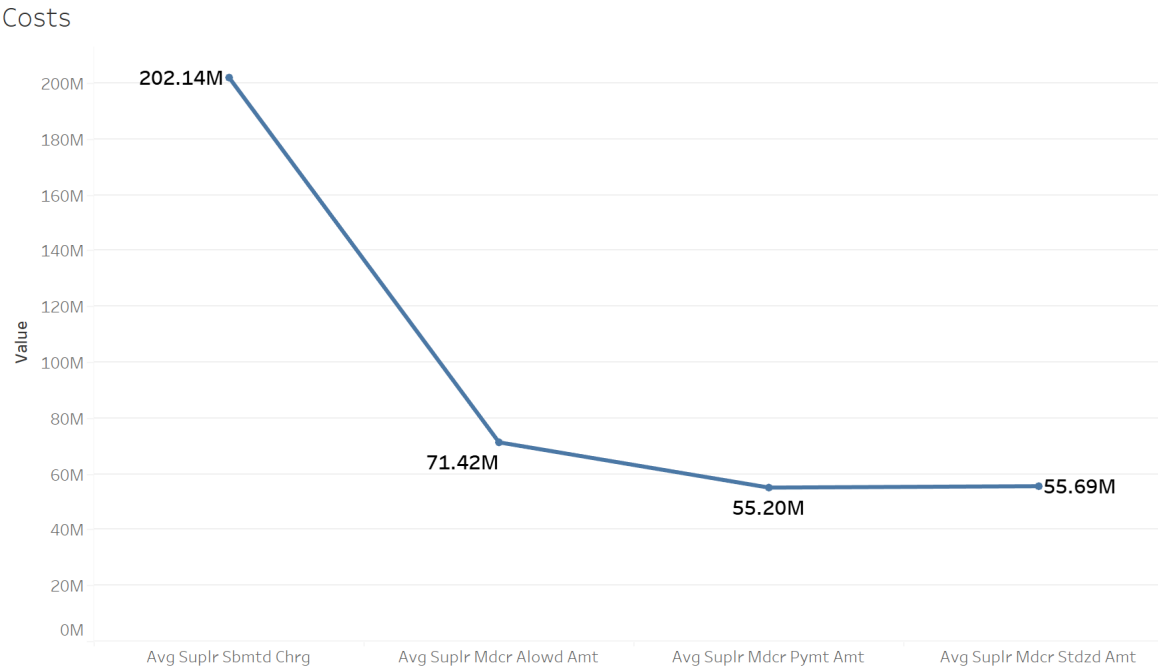
Statistical Analysis:

- This graph tells us whether the service offered by the supplier is either rented/ sold to the beneficiary. ‘Y’ for yes indicates that it is rented and ‘N’ for no indicates that the item is not eligible for rentals and is therefore usually sold.
- Under the **Durable Medical Equipment category**, almost half of them are eligible for rentals. This makes sense because beneficiaries only temporarily need equipment like

oxygen supplies and wheelchairs (usually for a few days following surgeries or temporary ailments) so it is better to offer them as rentals. This makes it affordable for the beneficiaries and allows repeated sales for suppliers as well.

- In the **Prosthetic and Orthotic Devices category**, no equipment is eligible for renting which is understandable because people almost always permanently need any form of prosthetic equipment they require for a lifetime.
- **Drugs and Nutritional Products** are almost always sold because they are consumed as medicine. The 2% rentals under this category belong to equipment associated with 'Enteral and parenteral' services. (Enteral nutrition refers to any method of feeding that uses the gastrointestinal (GI) tract to deliver nutrition and calories. Parenteral nutrition refers to the delivery of calories and nutrients into a vein from ACG)

Visualization 22: Medicare Coverage Analysis



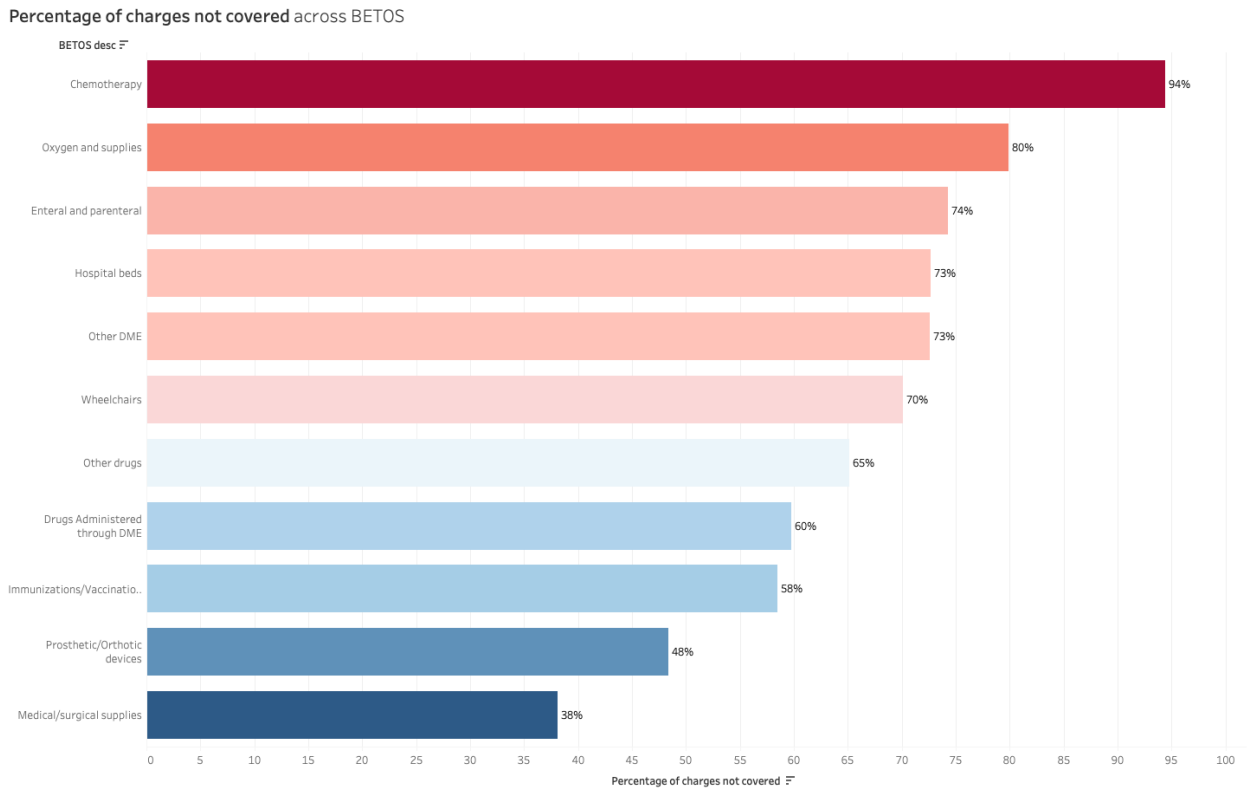
Statistical Analysis:

- Description of the charges specified in the above chart:
  - **Average Supplier Submitted Charge** is the amount charged by the supplier for providing a particular equipment or drug.
  - **Average Supplier Medicare Allowed Amount** is the average amount that Medicare allows for a particular service or item billed by the supplier.

- **Average Supplier Medicare Payment Amount** is the average amount that Medicare pays to the supplier for the service.
- **Average Supplier Medicare Standardized Amount** is the average standardized payment amount that Medicare would pay, adjusted to remove geographic differences in payment rates across different parts of the country.

- The supplier submitted charge (about 200M) is almost three times as high as what Medicare allows to pay for (71M). The actual payment amount is even lower than the allowed amount (55M) which is almost one-fourth of the supplier charge. This substantial disparity exposes a significant financial burden that may fall on patients or other insurance providers to cover the remaining costs.

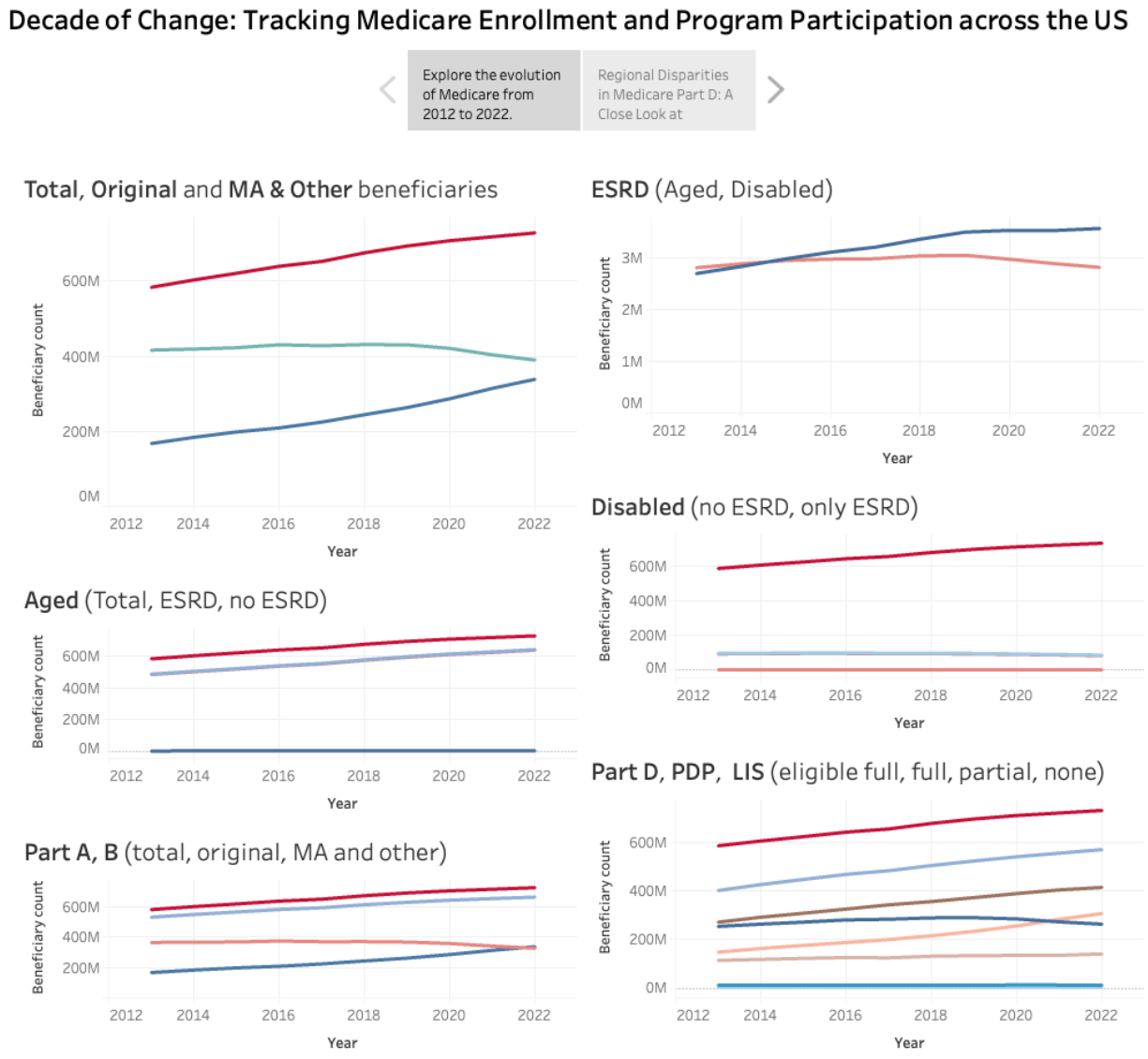
Visualization 23: Percentage of Costs not covered across BETOS categories



Statistical Analysis:

- **Chemotherapy** and **Oxygen and supplies** exhibit the highest non-payment rate higher than 75% indicating substantial coverage gaps for cancer patients and out-of-pocket expenses for essential respiratory aids.
- However, there is a considerable amount of coverage about 30% to 50% for supplier charges associated with **Wheelchairs** and **Prosthetic devices**.

Tableau Stories and Dashboards:

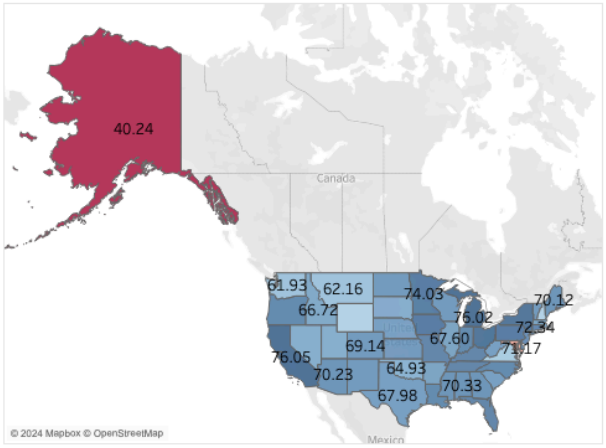


Decade of Change: Tracking Medicare Enrollment and Program Participation across the US

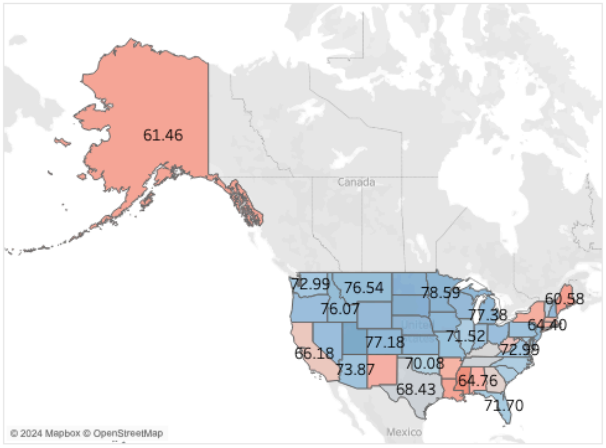
Explore the evolution of Medicare from 2012 to 2022.

Regional Disparities in Medicare Part D: A Close Look at

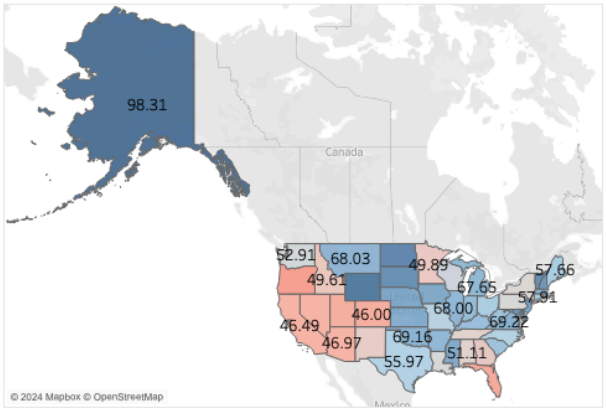
Part D in Total %



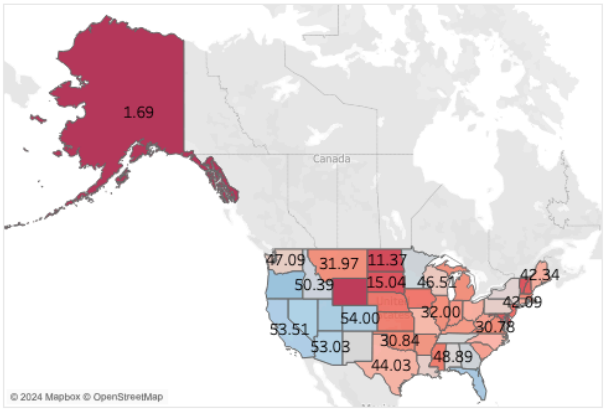
No LIS in Part D %



PDP in Part D %

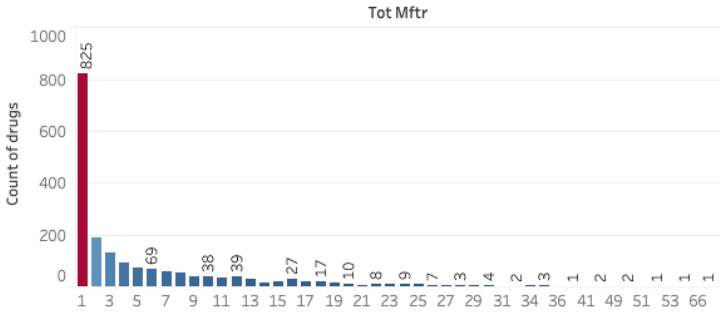


MAPD in Part D %

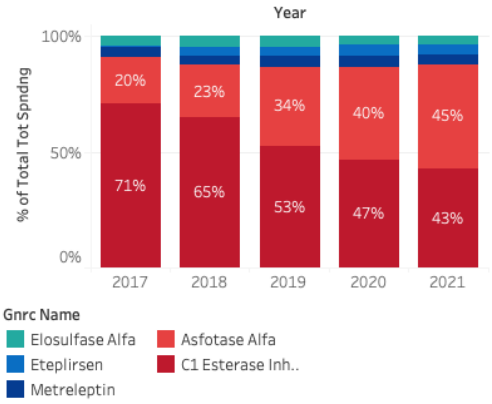


Medicare Drugs Analysis

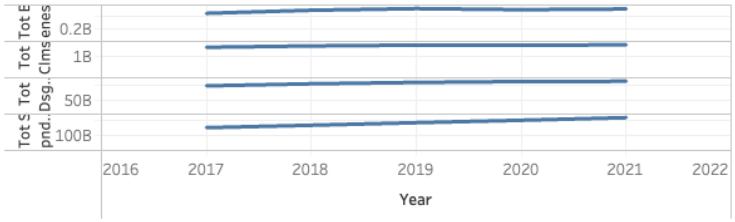
Drug count **by Manufacturer** count



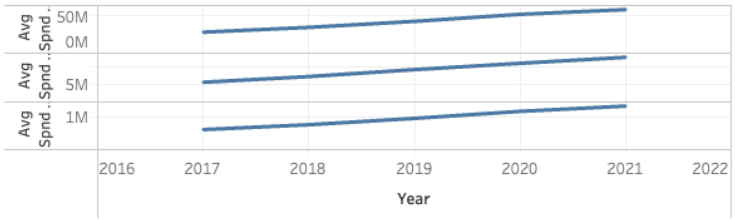
Popular drugs by spend over time



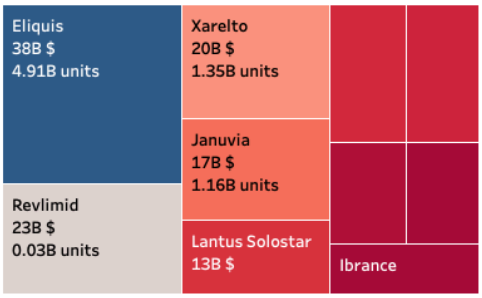
Total Beneficiaries, Claims, Dosage units, Spending



Average spend per Beneficiary, Claim, Dosage unit

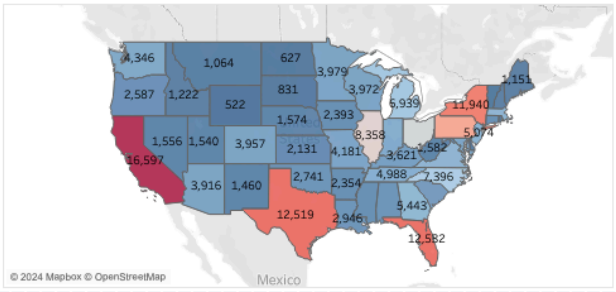


Top 10 brands by spend, dosage units

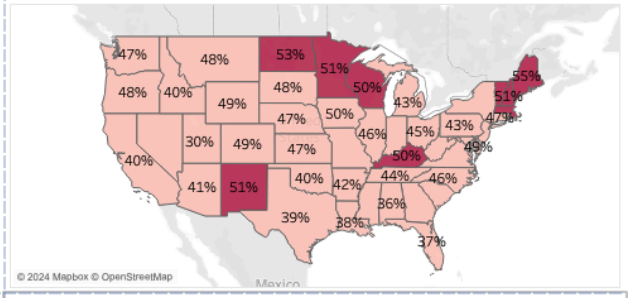


Medicare Physicians: Distribution, Demographics, and Practice Insights

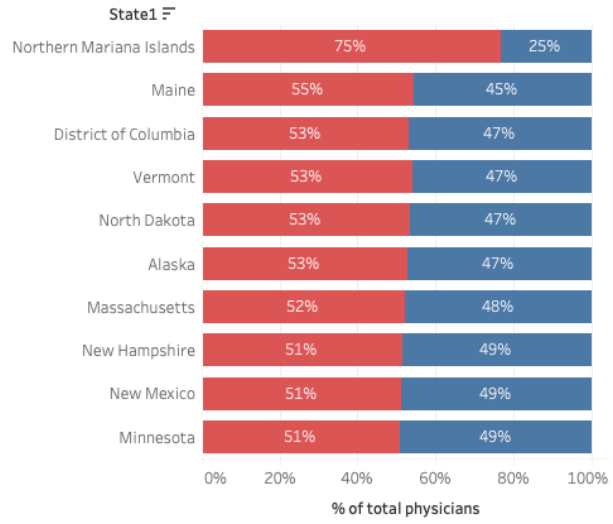
Physicians across the US



Female physicians across the US

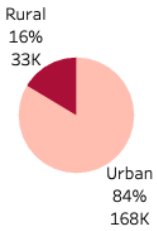


Top states by **Female physicians**

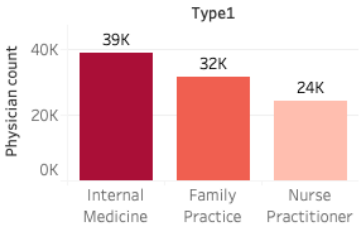


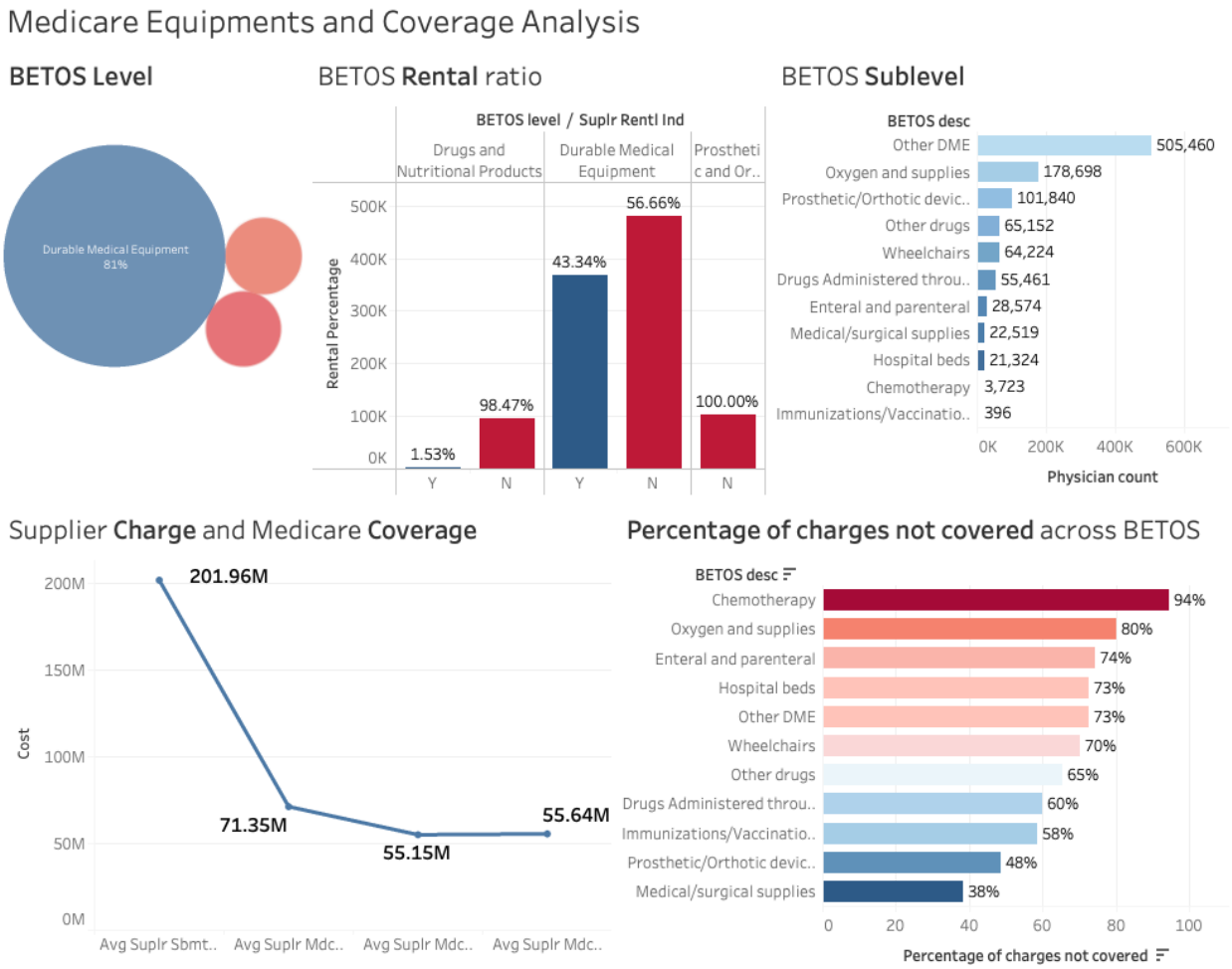
Physician **Credentials**

RUCA Category



Practice Types





Key findings:

<u>Medicare</u>

- Medicare Advantage and Other is popular over Original Medicare Plans in recent years and this pattern prevails in every analysis.
- The ratio of Aged and Disabled beneficiaries is approximately 4:1.
- Beneficiaries with ESRD are negligible.
- Almost 90% of the beneficiaries are enrolled in Part A, B, two-thirds in Part D and half do not qualify for LIS.

<u>Drugs</u>

- The majority of the drugs have fewer manufacturers and generic drugs have multiple manufacturers.
- Eliquis is the top manufacturing company by most spend and dosage units. However, the data cannot speak to its profitability due to a lack of information associated with production costs.
- C1 Esterase Inhibitor (prevents swelling from allergies), Asfotase Alfa (to treat the calcification of bones and teeth) are the most common drugs over time.
- Average spend per beneficiary, claim, and dosage unit grows with time. Total spending increases with time, and total claims and dosage units also gradually increase with time. Total beneficiaries have dropped in 2019 which could be attributed to the COVID-19 pandemic.

Physicians

- California, Texas, and New York have the highest number of physicians.
- Northern Marina Islands, Maine, and the District of Columbia have the highest percentage of female physicians.
- Almost 85% of physicians belong to the urban sector and the remaining 15% belong to the rural sector.
- Internal Medicine is the popular form of practice in the urban sector whereas in the rural sector, it is Family Medicine.
- Over 70% of Medicare physicians hold an MD medical degree
- Durable Medical Equipment (over 80%) is the prevalent type of BETOS equipment used and about 50% of it can be rented. The other two categories which are Prosthetics and Nutritional products (except Enteral and parenteral services) are not available for renting purposes.
- Medicare only covers about one-fourth of the charges submitted by the suppliers.
- Very little coverage is enabled for Chemotherapy and a considerable amount of coverage is allowed for Wheelchairs and Prosthetics.

Future Scope:

- The Association of Medicare and Home Health Agencies, a public or private organization that provides skilled nursing services to a patient's residence can be extensively analyzed and studied to understand their characteristics, trends, economic dynamics, and impacts on the well-being of beneficiaries.

- Interactive dynamics of enrollments and deaths (especially during the COVID-19 pandemic) can be examined.
- Profits specific to drugs and manufacturers can be interpreted if additional information on production costs were available.

Limitations:

Since a lot of the data are available in the form of multiple columns instead of subcategories of a single column, Tableau customizations to a certain extent felt limited. Since this was our first time learning and working with Tableau maybe we need more time to exercise all the functionalities that Tableau has to offer.

Lessons learned:

- We learned a lot during data understanding and transformation which was a huge task in the data preprocessing step. So far, we have only used Kaggle datasets for projects and they almost always come with data description and basic analysis to get you started. Since this data was taken directly from a data library, right from the beginning we had to be scrupulous in learning what each of the columns meant and how they can be transformed to get the kind of visualizations we have in mind.
- The extensive use of Tableau has left us feeling dexterous enough to handle standalone projects with it.
- From a personal perspective, we learned to stick with the data. Without anything to refer to, the project felt impossible as we felt like we did not even know where to begin with. But this project has made us believe in the virtue of persistence like nothing ever has before.

Conclusion:

The findings highlight the evolving landscape of Medicare. These could inform future policy adjustments aimed at addressing coverage gaps, controlling costs, and better tailoring services to meet the diverse needs of the Medicare population.